A Corpus-Based Study of Academic Vocabulary in Foundation-Level Students' Assessed Academic Writing at a UK University

Thesis

A corpus-based study of academic vocabulary in foundation-level students' assessed academic writing at a UK university

Dana Therova

(BA Hons, MA Hons, MRes)

Faculty of Wellbeing, Education and Language Studies

School of Languages and Applied Linguistics

Thesis submitted to The Open University for the degree of Doctor of Philosophy

February 2021

Abstract

Academic vocabulary is widely recognised as a key aspect of writing style in English for Academic Purposes (EAP) contexts, and there is a well-established link between the knowledge and use of academic vocabulary and academic achievement. Given that written assignments represent one of the main modes of assessment at UK universities, the significance of academic vocabulary in assessed academic writing seems indisputable.

This study reports on a textual analysis of the deployment of academic vocabulary in four genres of assessed academic writing produced by multilingual foundation-level students (N=193), in the context of one UK university in 2014 - 2018. It also investigates the deployment of academic vocabulary longitudinally by considering the development of these vocabulary items over the duration of a one-year foundation programme. Textual analysis is conducted through a corpus-based approach assisted by AntWordProfiler (Anthony, 2014) and AntConc (Anthony, 2018), which enable the identification and further exploration of academic vocabulary items on the basis of the New Academic Vocabulary List (Gardner & Davies, 2014). This is complemented by an online survey (N=51) and semi-structured interviews (N=14) exploring students' perceptions of the main factors contributing to the acquisition, deployment and development of academic vocabulary items in their written assignments.

Findings generated by textual analysis highlight the effects of the writing genre, topic and assignment brief on academic vocabulary in written production. The interview and survey findings underline the importance of an instructed environment and exposure to appropriate sources as well as the vital role of feedback, peer support and opportunities for practice. These findings have potentially important pedagogical implications for foundation programmes catering for diverse student populations.

Acknowledgements

I would like to express my sincere thanks to my supervisors Dr Maria Leedham and Dr Prithvi Shrestha for their invaluable advice and continuous support during my PhD study. I would also like to thank the late Professor Stephen Bax for his guidance during the initial stages of my research project. I also wish to express my gratitude to my examiners, Dr Paul Thompson and Dr Jackie Tuck, for their constructive comments and suggestions. I am also grateful to my participants for taking part in this study. Last but not least, I would like to thank The Open University for the opportunity and for granting me a studentship for full time PhD study.

Declaration of Authorship

I declare that this thesis has been composed solely by myself and that it has not been submitted, either in whole or in part, in any previous application for a degree. Except where otherwise acknowledged, the work presented is entirely my own.

Contents

Tables

Figures

Acronyms

AKL	Academic Keyword List
ANOVA	analysis of variance
AVL	New Academic Vocabulary List
AWL	Academic Word List
BAWE	British Academic Written English corpus
CI	confidence interval
CV	coefficient of variation
EAP	English for Academic Purposes
EGAP	English for General Academic Purposes
ESAP	English for Specific Academic Purposes
ESP	English for Specific Purposes
GSL	General Service List of English Words
HEI	higher education institution
IELTS	International English Language Testing System
IFP	International Foundation Programme
L1	first language / mother tongue / native language
L2	foreign language / second language/ additional language
M	mean average
MRU	Mauritius
NAWL	New Academic Word List
OPAL	Oxford Phrasal Academic List
RO	research oriented academic vocabulary
RO-desc	research oriented descriptive academic vocabulary
RO-loc	research oriented location academic vocabulary
RO-proc	research oriented procedure academic vocabulary
RO-top	research oriented topic academic vocabulary
RQ	research question
SD	standard deviation
SFL	Systemic Functional Linguistics
TO	text oriented academic vocabulary
TO-fram	text oriented framing academic vocabulary
TO-res	text oriented resultative academic vocabulary
TO-str	text oriented structuring academic vocabulary
TO-trans	text oriented transition academic vocabulary
UAE	United Arab Emirates
ZPD	zone of proximal development

Sub-corpora acronyms

E1_FN_113	Essay 1; Exposition Essay on ‘Fake news’
E1_LS_42	Essay 1; Exposition Essay on ‘Learning styles’
E1_MI_38	Essay 1; Exposition Essay on ‘Multiple intelligences’
E2_SMA_38	Essay 2; Discussion Essay on ‘Social media for academic purposes’
E2_SMC_42	Essay 2; Discussion Essay on ‘Social media and crime’
E2_SS_113	Essay 2; Discussion Essay on ‘Surveillance society’
PQ_BOS_48	Problem Question on ‘Business organisational structure’
PQ_ESP_73	Problem Question on ‘Ethics in social psychology’
PQ_SI_13	Problem Question on ‘Sustainable initiatives’
PQ_ST_59	Problem Question on ‘Sustainable tourism’
RR_193	Research Report on ‘Managing barriers to sustainability’

Chapter 1: Introduction

This study is set in the context of an international foundation programme (IFP) targeted at overseas students aspiring to pursue undergraduate study at a UK university. Specifically, since written assignments constitute the main form of assessment at UK universities (Lillis & Scott, 2007), the present study investigates foundation-level students' assessed academic writing with a particular focus on academic vocabulary, regarded as a specialised type of vocabulary prevalent in English for Academic Purposes (EAP) contexts (Coxhead, 2000; Gardner & Davies, 2014; Nation, 2001). This study has been motivated by my personal interest in university students' academic writing, which has been driven by my experience of working as an EAP practitioner in various higher education contexts in the UK, including IFP provision.

This chapter provides the background for the study by first describing the context of the higher education climate in the UK (1.1), followed by a discussion of the importance of academic writing in university settings (1.2). Next, the role of foundation-level courses at UK universities is focused on (1.3). This leads to the contributions of this study and research questions (1.4). The chapter concludes by outlining the organisation of this thesis (1.5).

1.1 Context of study

Due to internationalisation, the current UK higher education climate is characterised by a heterogeneous student body constituting both 'local' (i.e. British) as well as 'overseas' or 'international' (i.e. from outside the UK[1]) students (Lillis & Scott, 2007). In the academic year 2018 - 2019, for instance, there were 485,645 non-UK students in UK higher education institutions representing approximately 20% of the total student population. Out of these non-UK students, 268,395 (i.e. 11% of the total) were undergraduates, with 111,335 (i.e. 5%) enrolled in their first year (HESA, n.d.). This diversity in the social, ethnic and linguistic composition of student population inevitably leads to varying levels of preparedness for academic study in terms of the students' literacy skills, resulting in an urgent need to respond to students' learning needs and the

[1] While the terms 'overseas' and 'international' are both used to refer to students from outside the UK, a distinction is made in this study with 'overseas' describing all non-UK students and 'international' relating to students from outside the UK and EU, reflecting the Brexit situation i.e. at the time of writing the thesis (2020), the UK was in a period of transition, meaning that the same conditions applied to both EU and local students at UK universities.

demands that universities make of their students (Tribble & Wingate, 2013). The concept of 'academic literacy' is thus central in addressing students' needs, referring to "the ability to communicate competently in an academic discourse community" encompassing the skills of reading, presenting, debating, evaluating information and creating knowledge through speaking and writing, for instance, which need to be acquired by all students (local as well as overseas) new to an academic setting (Wingate, 2018, p. 350). This view of all students, regardless of their backgrounds, being novices to academic contexts is reflected in the Academic Literacies approach to student writing (Lea & Street, 1998), which has become an influential model in UK tertiary education. This approach is further discussed in 2.3.

In the context of UK universities, the specific academic literacy needs are catered for primarily by English for Academic Purposes (EAP) provision (further discussed in 2.2.2), which typically encompasses pre-sessional, in-sessional and foundation courses, each targeting different student populations. Pre-sessional courses are usually delivered over 4 - 12 weeks prior to the commencement of the students' undergraduate or postgraduate degree programmes and are targeted at international students who at the time of application do not possess the required level of language proficiency. In-sessional courses offer academic support primarily to international students during the academic year alongside the students' degree courses with the aim of mitigating students' lower language proficiency. Foundation courses are targeted at both local students who do not meet the university entry requirements in terms of their formal qualifications, as well as overseas students who do not meet the entry criteria required by UK universities in terms of their level of English proficiency and / or formal qualifications. These courses, which are typically 6 - 12 months long, aim to prepare students for undergraduate study and include 'undergraduate foundation programmes', 'university foundation courses', 'international foundation year courses', 'international year one courses' and equivalents (Home Office, 2019; Jordan, 2002; Pearson, 2020).

Despite being targeted at different student populations, the common goal of these EAP courses is to support the development of students' academic literacy skills. Out of all the academic literacy skills, it is academic writing that has often received most attention as the difficulties experienced by students at the level of academic literacy are often detected in their written production. Academic writing has thus often been a priority on the language agenda (Lillis & Scott, 2007; Wingate, 2018; Wingate & Tribble, 2012).

1.2 The role of academic writing in university contexts

The reason why writing has often received a great deal of attention on UK university EAP courses is the fact that writing is the main mode of demonstrating knowledge and understanding in university contexts, whereby written assignments constitute the principal means of assessment. Writing is hence regarded as a 'high stakes' activity in university education as students need to demonstrate the required standard of academic writing if they are to succeed in their university studies (Flowerdew, 2016; Lillis & Scott, 2007).

The importance of academic writing in turn emphasises two key aspects crucial for successful academic written production: the role of academic vocabulary (3.2.3) is considered a key element of academic writing style (Hyland & Tse, 2007) with insufficient knowledge of these lexical items being associated with a lack of academic success (Gardner & Davies, 2014); and genre awareness as students are assessed on the production of relevant genres in which they are required to express their knowledge (Wingate, 2019), with a 'genre' referring to "a term for grouping text together, representing how writers typically use language to respond to recurring situations" (Hyland, 2009, p. 15; further discussed in 2.2). It is, therefore, crucial for university students to develop awareness of the genres they are required to produce in their specific contexts (Wingate, 2018), including appropriate writing style incorporating some of the most common characteristics of academic writing, such as the deployment of specialised vocabulary prevalent in academic texts in order to follow the conventions placed upon them by the academic settings in which they operate.

Considering the highly prominent role that academic writing plays in university settings with different genres at the centre of written production and academic vocabulary widely recognised as a key feature of writing style in academic contexts, the importance of academic vocabulary in genres of assessed academic writing seems indisputable. It is, therefore, vital for students to not only acquire this type of vocabulary, but to also deploy this specialised vocabulary in their written production if they are to succeed in their studies. This is particularly important for students who are novices to academic settings and who possess very little or no prior experience with the type of writing required of them in university settings.

This study, therefore, explores the deployment and development of academic vocabulary in various genres of assessed academic writing produced by novice student writers in the context of one UK university. In particular, this study is set in the context of a foundation-level provision targeted at overseas students who aspire to pursue an undergraduate degree programme, but do not meet the entry requirements. These foundation courses play a crucial role in the UK higher education climate as they prepare students for study at a degree level by helping them develop the necessary academic literacy skills, discussed next.

1.3 Foundation-level provision at UK universities

With the increasing number of overseas as well as non-traditional[2] British students wishing to pursue university study, many universities offer a foundation-level course to prepare these students for academic study. The way in which foundation-level courses cater for these two groups of students vary, however, reflecting the distinction between 'local' and 'overseas' students (1.1). One type of foundation course is targeted at local students who do not meet the entry requirements for degree-level study in terms of their qualifications, do not feel ready for degree-level study, or are returning to study and wish to improve their academic skills to prepare for the demands of studying at a degree level. In addition, overseas students who meet the required language level, often at least 6.5 or 6.0 with no single element (i.e. reading, writing, listening, speaking) less than 6.0 or 5.5 respectively in the International English Language Testing System (IELTS) or equivalent, are also eligible to enrol on these foundation courses. While the UK was still an EU[3] member state, students from other EU countries could also apply to study on these courses, but did not have to provide proof of their English language proficiency.

The other group of students are overseas students (excluding those from EU[4] countries) with a lower language proficiency corresponding to a minimum of 4.5 IELTS or equivalent (with no single element less than 4.0). These students often enrol on foundation-level provision known as 'International Foundation Programmes' (IFPs).

[2] i.e. first-generation university attendees from ethnic minority backgrounds, working-class families and students with disabilities, who would not in previous generations have been expected to go to university, and whose limited knowledge of higher education means they may often experience difficulties in 'fitting in' at university (Christie, 2007; Holton, 2018)

[3] At the time of writing the thesis, the UK was in the Brexit transition period and it was not yet clear how or whether the entry requirements for EU students were going to change.

[4] The exclusion of EU students from the 'overseas' category of students is expected to change after the transition period (i.e. from January 2021).

Although there is no single structure followed by the IFPs delivered at the various UK higher education institutions, “one feature which is shared is the goal to prepare international students for university study” (Manning, 2008, p. 1). A further feature common to most IFPs is that on successful completion of the course students can progress onto a degree programme, usually within the same institution. In addition, IFPs also help orient students to university life by enabling them to familiarise themselves with the culture of UK higher education in general and of their selected university in particular.

A further distinction between these two types of foundation-level courses at UK universities is their focus on either a specific discipline or generic academic skills. Numerous foundation-level courses are discipline-specific, preparing students for their future degree study not only by developing their academic skills, but also by focusing on discipline-specific concepts, vocabulary and written genres. In contrast, discipline non-specific foundation-level provision, such as generic IFPs, focus primarily on the development of the students’ language and academic skills.

This study is set in the context of one such generic IFP offered by a British university (henceforth ‘University’) based in the South-East of England and its two overseas campuses located in Mauritius and the United Arab Emirates (UAE). The IFP in question runs over one academic year, equalling 24 teaching weeks delivered over six months. Hence, there is a lot of pressure to develop the students’ skills to a required level in a relatively short time. The selected IFP focuses on the development of the skills deemed necessary for degree study. Specifically, the IFP in this study aims to provide a structured framework to support students in becoming self-directed learners in acquiring a wide range of transferable skills essential for undergraduate study. The selected IFP hence meets the general aims of foundation-level courses at UK universities to prepare international students for study at a degree level. This is achieved by helping students to improve their general level of English and develop academic skills, such as independent learning, academic writing, researching and presentation skills, regarded as important at British universities. These skills are assessed through several summative tasks throughout the IFP, meaning that students do not progress automatically to a degree programme on completion of the IFP. The summative assignments include both individual as well as group assessments in the form of predominantly written assignments with fewer oral assessment tasks (4.3.4). On successful completion of the

IFP, the students are able to progress to a degree programme of their choice within the University as they have demonstrated that they have acquired a range of transferable skills necessary for undergraduate achievement within the UK higher education system (University's IFP prospectus, 2014; University's IFP specification, 2014).

Since IFPs are targeted at overseas students, the programmes are characterised by a diverse student population. In this study, the participants (N=193) are predominantly young adults aged 16 - 26 (M=19) representing 54 nationalities and 55 linguistic backgrounds (Appendix 4). The variety of language backgrounds represented in the current study is suitable for the investigation of intergroup homogeneity, referring to similarities among L2 writers irrespective of their linguistic backgrounds, which can provide insights into language features inclusive to L2 writers as a composite. Such an approach can be considered important as it can provide evidence that some aspects of L2 writing may be related to the students' language experience or level of proficiency rather than being independent or cultural (Crossley & McNamara, 2011). The intergroup homogeneity approach to the investigation of student writing is relevant to this study due to the participants' similar level of proficiency (i.e. for entry to this IFP the students must have achieved a minimum language level of IELTS of 5.5 with a minimum of IELTS 5.0 in each component).

Considering the importance of writing in university contexts where written assignments are the main way of assessment (1.2), the vital role of effective writing skills is indisputable. Academic writing produced by university students has thus been investigated extensively through different approaches and focusing on various aspects of student writing, outlined next.

1.4 Contribution of study and research questions

An extensive body of research exists in the domain of university student writing. Some of these studies examine the similarities and differences between writing produced by student and expert writers (e.g., Durrant & Mathews-Aydinli, 2011; Flowerdew, 2003; Qin, 2014). Numerous studies compare student writing across various academic disciplines (e.g., Durrant, 2014; Hyland & Tse, 2007) or investigate student writing within a specific academic domain (e.g., Baratta, 2010; Mei & Allison, 2005; Shrestha, 2020; Shrestha & Coffin, 2012). Extensive research has also been conducted in the area of student writing produced by speakers of English as their first language (L1) and second

language (L2), predominantly investigating the similarities and differences between L1 and L2 speakers' written production (e.g., Ädel & Erman, 2012; Hinkel, 2003; Staples & Reppen, 2016). A considerable number of studies also focus on linguistically homogeneous groups of students; that is, writing produced by students from the same linguistic background (e.g., Huang, 2015; Leedham, 2012; Li & Schmitt, 2010; Plakans & Gebril, 2012; Zhang, 2011). Several studies have taken a longitudinal approach to the investigation of various aspects of student academic writing (e.g., Knoch et al., 2015; 2014; Leedham, 2006; Oppenheimer et al., 2017; Serrano et al., 2012) or focused on various genres of student writing (e.g., Nesi & Gardner, 2012; Olinghouse & Wilson, 2013).

However, despite the important role that foundation-level courses play in the context of UK higher education (1.3), the writing of foundation-level students in the context of UK universities remains under-researched. This study hence aims to address this significant omission, and contribute to the current body of literature on student writing by investigating the assessed academic writing produced by a multilingual group of foundation-level students with a particular focus on academic vocabulary items across different genres of assessed academic writing. In attempting to fill this gap, this study addresses the following research questions (RQs):

1. What are the density and diversity of academic vocabulary in international foundation students' assessed academic writing across writing genres?

2. To what extent do the density and diversity of academic vocabulary in international foundation students' assessed academic writing develop over an academic year?

3. What are the students' perceptions of the contributing factors impacting the acquisition, deployment and development of their academic vocabulary?

RQ1 and RQ2 are based on textual analysis of the usage of academic vocabulary items in terms of coverage (i.e. density) and variety (i.e. diversity) of these vocabulary items in the students' texts, primarily from a quantitative perspective, while RQ3 provides the learners' perspectives on their experience with academic vocabulary on the IFP with a more qualitative focus. It should, therefore, be noted that the present study does not consider the deployment of academic vocabulary from the perspective of appropriate usage.

Addressing the above research questions is intended to broaden the understanding of multilingual foundation-level students' usage of academic vocabulary in written production, thereby informing foundation programmes as well as broader EAP provision offering academic writing support to a diverse student population.

1.5 Thesis overview

Chapter 1 has set the context of this study by outlining the current higher education climate in the UK. It has discussed the importance of academic writing skills in university settings, which in turn highlighted the key role of genre awareness and academic vocabulary. This chapter has also focused on the role that foundation-level provision plays at UK universities and has indicated the gap in the current body of research, which the present study aims to address. Following the Introduction, this thesis contains seven further chapters.

Chapter 2 focuses on approaches commonly taken to the investigation of student writing. These include various genre-based approaches including Systemic Functional Linguistics (SFL), English for Specific Purposes (ESP) and Corpus Linguistics, as well as Academic Literacies. The discussion of these approaches also includes a review of the relevant literature, further emphasising the omission of IFP courses in the current body of research.

Chapter 3 is concerned with vocabulary research. It first explains what counts as a word and outlines the different types of vocabulary, then provides an overview of what is involved in vocabulary knowledge and learning. This leads to a discussion of academic vocabulary as a specialised type of vocabulary prevalent in academic contexts and of the various academic word lists presently available. It also provides an overview of academic vocabulary research drawing on academic word lists and outlines vocabulary research relating to foundation-level students, which highlights the gap in academic vocabulary research in the context of IFPs.

Chapter 4 outlines the theoretical frameworks and approaches to the investigation of writing and how these informed the data collection procedure in this study. In this chapter, the research context is described in greater detail together with the recruitment procedures, ethical considerations and an overview of the participants and the collected data. The selected academic word list forming the basis of this study is also

justified and a detailed account of the data analysis in relation to individual research questions (RQs) is given.

The next three chapters present and discuss the findings relating to individual RQs. Chapter 5 relates to RQ1 and presents the findings and discussion concerning academic vocabulary usage across the different writing genres of the foundation-level students' assessed academic writing investigated in this study. Chapter 6 addresses RQ2, which focuses on academic vocabulary from a longitudinal perspective by exploring the changes in the deployment of academic vocabulary in the participants' writing over the duration of the foundation programme under study. Chapter 7 reports on the findings relating to RQ3 and provides a discussion of the students' perceptions of the main factors that impacted the acquisition, usage and development of their academic vocabulary.

Chapter 8 concludes the thesis by first summarising the study's main findings. This is followed by the contribution that the current study makes to research, methodology and pedagogy. Next, it outlines the limitations of this study, which leads to recommendations for future research.

Chapter 2: Approaches to the investigation of student writing

2.1 Introduction

Having established the vital role that academic writing plays in university students' academic achievement (1.2), this chapter discusses the various approaches to the investigation of writing frequently applied to the examination of student writing in university contexts. Specifically, the focus of this chapter is on genre-based perspectives (2.2) encompassing Systemic Functional Linguistics (2.2.1), English for Specific Purposes (2.2.2) and Corpus Linguistics (2.2.3), followed by a discussion of how these approaches can be used in a complementary manner, including for classification of genres of university student writing (2.2.4). Next, the Academic Literacies approach is discussed (2.3) due to its influence on academic writing research. The chapter concludes by summarising these approaches, pointing to a gap in the current body of literature which this study attempts to address (2.4).

2.2 Genre-based approaches

A 'genre' is a term for grouping texts together based on how language is typically used to respond to recurring contexts and situations, with each genre displaying a number of characteristics, such as its particular purpose, overall structure, or specific linguistic features, which differentiate it from other genres (Hyland, 2009). The concept of a genre is underpinned by the idea that texts of the same kind are easily recognisable by the members of the community who frequently use these texts as they have certain expectations of them based on their previous experiences with texts of the same kind and are hence able to draw on their repeated experiences with these texts. This relates to two assumptions underpinning genre analysis: the first assumption concerns the fact that features of a similar kind of texts are dependent on the social context in which the texts were produced; the second assumption is based on the premise that it is possible to describe text features in a way that relates a text to similar texts (Hyland, 2002, 2004, 2009).

There are various traditions exploring the concept of a genre that have informed current understandings and applications of genre, and several taxonomies for dichotomising genre theories have been proposed by researchers in the field. Hyon (1996), for instance, divides genre theories into three research areas comprising English for Specific Purposes, North American New Rhetoric studies, and Australian Systemic Functional

Linguistics, while Flowerdew (2002) categorises genre theories into linguistic and non-linguistic camps with the former encompassing the Australian school and English for Specific Purposes and the latter represented by the New Rhetoric. More recently, Bawarshi and Reiff (2010) propose a classification which includes literary traditions (e.g. Neoclassical, Structuralist, Romantic or Post-romantic approaches to genre), rhetorical and sociological traditions (e.g. the New Rhetoric or Social Phenomenology) and linguistic traditions (e.g. Systemic Functional Linguistics, English for Specific Purposes and Corpus Linguistics).

Due to its focus on writing research in EAP contexts (seen as a branch of ESP) and use of the methodology of corpus linguistics, the current study follows Bawarshi and Reiff's (2010) genre taxonomy with the linguistic tradition being of particular relevance to this study. It is noteworthy, however, that while these approaches provide valuable means of researching genres in various contexts, the scholars within the linguistic tradition understand the concept of a 'genre' in different ways. The genre approaches representing the linguistic tradition are, therefore, focused on in the next sections (2.2.1 - 2.2.3). This is followed by a discussion of how these genre approaches can be combined, leading to categorisation of genres of student academic writing (2.2.4) with the aim of informing the approaches to the investigation of student writing and the genre categorisation adopted in this study.

2.2.1 Systemic Functional Linguistics

Based on the work of Halliday (1985) and applied to genre particularly in the work of Martin (1997), Systemic Functional Linguistics (SFL), has made a significant contribution to both genre research and pedagogy. Within this tradition, 'systemic' is used to refer to the structure or organisation of language (i.e. systems of choices) available to language users, while 'functional' refers to the work or function that language has in specific contexts (Bawarshi & Reiff, 2010; Eggins, 2004). Hence, as meaning resides in systemic patterns of choice, a language is a resource for making meaning (Halliday, 2014). Accordingly, language is a system of choices which enables writers to communicate certain functions, thereby allowing them to interact with others, express their experiences of the world and create coherent messages (Hyland, 2004).

In SFL, *genre* is defined as a "system of staged, goal-oriented social processes through which social subjects in a given culture live their lives" (Martin, 1997, p. 13), emphasising

the interactive nature of genres and the way they are linked to their context. According to Martin's definition, genres are staged because meanings are created in steps and it usually takes more than one step to work through a genre; genres are goal-oriented because they aim to achieve something; and they are social processes because genres are undertaken interactively with others (Martin, 2009). The emphasis in SFL is, therefore, on the social purposes of genres (also referred to as *text-types* in the SFL tradition) together with the rhetorical structures serving these purposes.

Genre analysis in SFL rests on the premise that when a number of texts share a purpose, they likely display the same structure and hence belong to the same genre. The relationship between texts and contexts is thus regarded as central in SFL as texts can only be understood in their social settings, with each context having the potential for a variety of possible texts (Hyland, 2004). SFL considers the relationship between texts and contexts at the level of *register* and *genre*. *Register* functions at the level of context of situation, where meaning is projected onto social context; *genre* functions at the level of culture by specifying how a given culture organises this meaning through stages in each genre. In the process of text creation, choices are first made at the level of context of situation regarding the *register* along the following three dimensions: *field* referring to the social activity including descriptions of the participants, subject matter and circumstances that these activities involve; *tenor* which is concerned with the relationship of the participants in the activity; and *mode* referring to the role of the language in the activity (Halliday, 2014; Martin, 1997, 2009, 2016).

These three dimensions at the level of context correspond to three language metafunctions at the linguistic level: *ideational* (concerned with representation; a resource for naturalising reality and building field knowledge, thereby enabling participation in activities), *interpersonal* (concerned with interaction; a resource for valuing these activities and negotiation of social relations, enacting tenor), and *textual* (concerned with information flow; a resource for managing the communicative demands of discourse by phasing ideational and interpersonal meaning together in mode) (Martin, 1997, 2009, 2016).

These dimensions impact the language choice of the text writers, resulting in some registers displaying predictable lexical and grammatical features. At the level of *genre* relating to the context of culture, the writer's linguistic choices are influenced by their social purpose in using language. *Genre* is thus realised through *register* (encompassing

field, mode and tenor) and involves certain conventions for text organisation so that the text's purpose can be recognised by readers (Hyland, 2004; Martin, 2016).

The relationship between *register* and *genre* provides an understanding of how social purposes are connected to text structures and how they manifest themselves within register. This move from the identification of social purpose to the analysis of a text's register (represented in field, tenor and mode) to linguistic analysis (involving semantic, lexico-grammatical or phonological features) is regarded as the most common trajectory in SFL genre analysis (Bawarshi & Reiff, 2010). This trajectory for the analysis of writing has shown that different types of texts display distinctive patterns of grammar, vocabulary and cohesion, which structure the texts into stages with each stage supporting the purpose of the genre. This makes it possible in SFL to describe texts in terms of the functions they serve as well as in terms of how these functions are expressed by the component linguistic elements of the texts. Defining these genre components linguistically enables genre analysts to explore how these elements combine in various ways to form the genres common in various contexts.

Several core educational genres have been identified by SFL theorists, initially in the context of primary and secondary school texts and later in adult migrant English education and workplace training programmes (Martin, 2009). These genres include: *exposition* (giving arguments for a thesis), *explanation* (giving reasons), *description* (giving an account of events), *narrative* (providing a reflection on experience), *report* (presenting factual information), *procedure* (showing how to do something) or *recount* (reconstructing past experiences) (Hyland, 2004; Hyon, 1996). It is also noteworthy that the term *macro-genre* is often used to refer to more complex genres encompassing other *elemental genres* (Hyland, 2002, 2004).

In terms of pedagogy and teaching the writing of these genres, SFL has drawn on the work of Vygotsky (1978), who put great emphasis on learning as a result of a verbal interaction and task negotiation with a more knowledgeable person (e.g. a teacher or a more capable peer), who has a central role in so-called *scaffolding* referring to the process of guiding student development through a number of cyclical stages including contextualising the genre, modelling appropriate rhetorical patterns of the genre, providing guided practice in writing and finally withdrawing to enable learners to write independently. These processes are closely related to the notion of the *Zone of Proximal Development (ZPD)*, used to describe the area between what learners can do

independently and what they can do with assistance; that is, the gap between the learners' existing level of performance and a level of potential performance, which can be bridged through social interaction and the assistance of more knowledgeable others, whereby seeing learning as a type of cognitive apprenticeship (Lantolf, 2000).

In accordance with Vygotsky's views of learning, SFL pedagogy has developed a model of learning, referred to as the teaching / learning cycle. The cycle involves working with texts with gradually increasing complexity, starting with a 'deconstruction' stage during which models of the target genre are presented, followed by a 'joint construction' stage where the learners and teachers collaboratively construct a text on the basis of the model that emerged from the deconstruction stage, leading to an 'independent construction' stage where the learners produce a text modelled on those from the preceding stages (Martin, 2009).

Even though SFL-informed pedagogy has been successfully applied in various teaching contexts, some have criticised the complexities of SFL for teaching purposes, however. Bourke (2005, p. 93) notes that in the context of L2 teaching the SFL approach is too complex due to its "grid of system networks", where meaning is processed simultaneously on three levels (i.e. ideational, interpersonal, textual) with the output from each stage becoming the input to the next. According to Bourke (2005), hence, SFL pedagogy lacks one of the key characteristics necessary for pedagogy; that is, simplicity. In the context of teaching academic literacy to L2 learners, further limitations of SFL-based pedagogy relate to the approach being too technical to be a viable framework for pedagogical purposes and potentially resulting in a static reproduction of various genres rather than in a critical analysis of the genres required in various disciplines (Gebhard, 2010). Further, the focus of SFL pedagogy on text reproduction often leads to students producing poor attempts at the target writing genres due to the limitations placed on students' understanding of texts (Knapp & Watkins, 2005). These criticisms of SFL pedagogy have been addressed by Rose and Martin (2012), who have attempted to simplify the SFL approach for teachers and students.

Numerous researchers have adopted the SFL framework to the investigation of university student writing. In the context of one university in India, Ndoricimpa (2019) draws on SFL to investigate 67 postgraduate students' texts from the perspective of the ideational, interpersonal and textual metafunctions at the sentence, paragraph and text levels. His findings reveal that students encounter difficulties relating to all three

metafunctions. Ndoricimpa (2019) thus calls for practitioners to explicitly focus on how the three metafunctions are achieved in the essay genre investigated in his study and also emphasises the importance of giving students feedback and support in writing skills, grammar and other relevant language skills. The textual, interpersonal, and ideational resources in essays were also explored in the context of the University of California by analysing two expository essays produced by two L2 learners, one awarded a fail mark and one awarded a pass mark (de Oliveira, 2015). This study found different use of the resources by the less successful and the more successful writer. Specifically, the essay awarded a pass mark displayed a more detached style compared to the more personal style in the essay that received a fail mark.

Similarly, less competent and more competent student writing (based on the researcher's evaluation of topic contextualisation, coherence, cohesion and thematic development) was investigated in the context of first year undergraduate students at a University in South Africa with the aim of exploring how the two groups of students develop themes in their writing (Tshotsho, 2014). The results show that the group of less competent students were less able to write coherent texts than their more competent peers as they failed to contextualise the topic and to link their sentences and paragraphs. Tshotsho (2014) attributes this lack of coherence displayed in the students' texts primarily to insufficient exposure to reading materials and argues that it is vital for students to be guided through the essay writing process by their teachers. Next, in the context of one Australian university, analysis of undergraduate students' writing with the aim of exploring how their written assignments contribute to the students' disciplinary knowledge was conducted by Woodward-Kron (2005). She found that the most common assignments in the field of education were expositions and discussions and that these macro-genres contained characteristics of various other genres embedded as micro-genres (i.e. genres embedded in the macro-genre), which help students develop their disciplinary knowledge. Woodward-Kron (2005) thus emphasises the importance of EAP practitioners' understanding of discourse structure and variation in disciplinary contexts.

These studies demonstrate the application and contribution of SFL to the investigation and teaching of the different aspects of student writing in various university contexts. Nevertheless, some of the criticisms levelled at SFL-based research include its focus predominantly on writing as a product as traditionally SFL linguists focus primarily on

the language outcomes as a result of the context and the social activities in which people engage (Prosser & Webb, 1994). A further critique is that due to the focus on cohesion, other important linguistic features of texts, such as collocations, have been neglected in SFL research (Martin, 2016). In addition, some of the studies drawing on SFL (e.g., de Oliveira, 2015; Tshotsho, 2014) tend to compare the writing of successful and less successful student writing, potentially treating the less proficient writing as deficit.

In sum, the SFL approach is characterised by its focus on texts in context and its emphasis on the relationship between language, text and context with the aim of making this relationship explicit, thereby providing insights into the relationship between language and society. Although the text is the primary unit of analysis in this approach, SFL text analysis goes beyond the analysis of linguistic resources and encompasses the analysis of the social, cultural and ideological meanings surrounding these texts. Hence, in SFL, text analysis cannot be separated from the context of use.

2.2.2 English for Specific Purposes

English for Specific Purposes (ESP), associated primarily with the work of Swales (1990), and further developed by Bhatia (2004), is primarily interested in genre as a tool for investigating and teaching the kinds of writing required of L2 speakers of English operating in professional and academic contexts, stemming from the demands placed on competent communication seen as fundamental in these contexts. Genre analysis in ESP is thus defined as the study of situated linguistic behaviour in institutionalised academic or professional settings, where ESP is often used as a hypernym encompassing more specialised areas of study such as English for Academic Purposes (EAP), English for Occupational Purposes (EOP) or English for Professional Communication (EPC) (Bawarshi & Reiff, 2010; Bhatia, 1995; Hyon, 1996).

According to Swales (1990, p. 58) "[a] genre comprises a class of communicative events, the members of which share some set of communicative purposes". These purposes constitute the rationale for the genre as they are recognised by the expert members of the discourse community. ESP thus adopts a somewhat narrower concept of genre than SFL by regarding genres as the property of specific discourse communities rather than seeing them related to the wider culture. The concept of a *discourse community* is hence central to ESP, and is defined as having the following: an agreed set of common goals, mechanisms of intercommunication among its members, mechanisms for providing

information and feedback, community specific genres, specialised terminology and a high level of relevant content and discoursal expertise (Swales, 1990).

The ESP perspective of a genre therefore places emphasis on the communicative needs of particular discourse communities with the aim of helping students to master the functions, stylistic and organisational features and linguistic conventions of the texts they are required to read and produce in their professions and disciplines. Genres in ESP are thus seen as the social actions commonly used and recognised by community members who use these genres in a particular context to achieve a particular purpose written for a particular audience (Hyland, 2004; Hyon, 1996).

Some of the most common ESP genres in the academic contexts include research articles, conference abstracts, book reviews, grant proposals, undergraduate essays, PhD theses and textbooks. In professional contexts, ESP genres encompass engineering reports, legal cases and briefs, email memos, company annual reports, charity donation requests, letters of recommendation, business faxes or emails, for instance (Hyland, 2004). Despite this wide range of genres, ESP genre analysis is commonly associated with Swales's (1990) move analysis as exemplified through research article introductions with a CARS (i.e. Create A Research Space) model comprising the following moves: 1) establishing a territory (by claiming centrality and/or making topic generalisations and/or reviewing previous research); 2) establishing a niche (by counter-claiming or indicating a gap or question-raising or continuing a tradition); 3) occupying the niche (by outlining purposes or announcing present research, by announcing principal findings and by indicating the structure of the research article) (Swales, 1990).

The CARS model thus describes the rhetorical patterning of a specific genre, with each move representing a distinctive communicative act intended to achieve one primary communicative function. This kind of analysis hence involves the identification of a series of moves and steps within a move forming the genre on the basis of a representative sample of texts. That is, owing to the communicative purpose defined in relation to a discourse community's shared goal, an ESP approach to genre analysis typically begins by identification of a genre within a discourse community together with defining the genre's communicative purpose. This is followed by an examination of the genre's organisation (i.e. its schematic structure often characterised by rhetorical moves) and an examination of the textual and linguistic features such as voice, grammar, syntax or style. The typical trajectory in ESP genre analysis therefore proceeds from

context to textual analysis with various levels of linguistic analyses applied at the textual level; that is, ESP genre analysis moves from identifying the genre's purpose to analysing the genre's rhetorical moves and how these moves manifest themselves textually and linguistically (Bawarshi & Reiff, 2010).

Another example of ESP genre analysis providing insights into the rhetorical patterning of a specific genre is Swales and Feak's (2010) analysis of research article abstracts, who identified the following characteristics of research article abstracts: 1) background / introduction / situation; 2) present research / purpose; 3) methods / materials / subjects / procedures 4) results / findings; 5) discussion / conclusion / implications / recommendations. Similarly, the moves in conference abstracts have been explored with the following identified moves: 1) outlining the research field; 2) justifying a particular research / study; 3) introducing the paper to be presented; 4) summarising the paper; 5) highlighting its outcomes / results (Yakhontova, 2002).

The genre analysis within the ESP tradition has therefore made invaluable contributions to the knowledge and description of discipline-specific genres, most notably research articles and other closely related genres, such as conference abstracts. Knowledge of these genres is regarded as important, particularly for graduate-level L2 speakers of English, as it has enabled them to not only gain access to academic and professional discourse communities, but also to actively participate in them (Bawarshi & Reiff, 2010).

The approach to ESP genre analysis is closely reflected in ESP pedagogy, catering primarily for adult L2 speakers of English with specific communicative needs in tertiary education or professional settings. In an ESP approach to pedagogy, in addition to the central notions of *genre* and *discourse community*, the notion of *task* is also emphasised, referring to the procedures involved in encoding and decoding of genre-related aspects of text-role and text-environment defined as "a set of differentiated, sequenceable goal-directed activities" (Swales, 1990, p. 76). These three key elements are inter-related as genres are considered the property of discourse communities, and the procedures involved with the processing of genres can be regarded as tasks; and genres, task and discourse community are linked together by communicative purpose which all of them share (Flowerdew, 2015).

Since many students may encounter difficulties with the genres they are exposed to, particularly with regard to organisation and phraseology, the overarching aim of ESP pedagogy is to facilitate students' transition into becoming an accepted member of their

discourse communities. The ESP approach to pedagogy is characterised by achieving this by putting great emphasis on rhetorical consciousness raising, seen as a fundamental feature of the ESP approach (Swales, 1990), referring to the ways in which instructors guide learners in developing an understanding of genres and communities. In a classroom, this is accomplished by means of analyses of the genres required of students with a focus on rhetorical action as well as organisational and linguistic means of its accomplishment while considering the audience (i.e. discourse community) (Flowerdew, 2015; Hyland, 2004). ESP pedagogy can thus assist students in examining not only the similarities within a genre, but can also raise their awareness of the differences in communicative event as well as individual language users (Devitt, 2015). Through genre analysis, students are hence made aware of the recurrent textual patterns, which allows them to participate in the activities of their target discourse communities.

Although Swales does not theorise his pedagogical approach to genre in Vygotskyan (1978) terms, numerous features of Vygotsky's (1978) approach can be found in ESP (Flowerdew, 2015; Hyland, 2004). These relate to the concept of 'consciousness raising' including Swales's definition of a task being 'sequenceable' and teachers guiding students through the various stages of the assigned tasks, which could be regarded as corresponding to the apprenticeship approach with a teacher scaffolding the students' work, meaning that scaffolding is an important aspect of genre-based ESP pedagogies. However, Dudley-Evans (1995) points out that although there is much value in an ESP approach to the teaching of academic writing for both experienced and inexperienced writers, there are also drawbacks in an over-prescriptive approach when focussing on genres in the classroom, relating mainly to implying that all writers need to do is learn a basic move structure for each section of the research article.

In sum, the ESP approach views genres in relation to the communities in which they are used and aims to provide insights into the rhetorical complexities of texts as a function of the choices and constrains placed on text producers when engaging in community practices (Hyland, 2015). This approach hence combines rhetoric and linguistics to study genre as grounded in shared communicative purposes, realised in linguistic utterances and discoverable through text analysis by making links between rhetorical moves, linguistic patterns, communicative purposes and discourse community (Devitt, 2015) with the goal to assist learners in becoming more successful readers and writers of the sort of texts they encounter in academic and workplace settings (Hyon, 1996). In

academic settings, where writing is the main mode of expressing disciplinary knowledge and understanding (1.2), students entering universities are often uncertain about the expectations of studying at a university level, particularly in terms of the type of writing they are required to do in their disciplines. This has been addressed by English for Academic Purposes (EAP), regarded as a branch of ESP, discussed next.

English for Academic Purposes (EAP)

EAP in university contexts has developed in response to the demands placed on students who must acquire the conventions of the type of language used in academic settings to successfully navigate their learning. The aim of EAP in university settings is hence to equip students with the skills they need in order to participate in specific academic contexts by providing insights into the structures and meanings of academic texts as well as into the demands that academic contexts place on communicative behaviours (Hyland & Hamp-Lyons, 2002).

EAP study is thus characterised by its focus on specific, purposeful uses of language reflecting a specific practical need in specific contexts, reflecting its origins in the wider English for Specific Purposes (ESP) contexts, with a common distinction frequently made between English for General Academic Purposes (EGAP) and English for Specific Academic Purposes (ESAP) within EAP provision. The former focuses on the skills, language and study activities thought to be shared by academic disciplines, based on the assumption that there are generic academic skills that are transferable across academic contexts and once students have mastered these generic features they can utilise them in a variety of academic contexts. The latter is concerned with the rhetorical and linguistic demands associated with particular academic contexts and disciplines. Despite the division into two branches, the overarching goal of EAP, common to both EGAP and ESAP, is to equip students with the knowledge and skills necessary to succeed in their studies (Flowerdew & Peacock, 2001; Hyland, 2016a).

Hence, some of the main characteristics of EAP include its content being related to particular activities together with a focus on the language appropriate to those activities in terms of lexis, discourse, syntax or semantics, for instance (Strevens 1988 in Flowerdew & Peacock, 2001). EAP pedagogy thus offers insights into this special type of language commonly used in academic settings and encompasses instruction focusing on the needs and practices of the kind of literacy required of L2 learners in academic

contexts so as to help them become familiar with the constraints that academic contexts place on language use (Hyland & Shaw, 2016).

This is closely related to the main principles on which EAP rests, as proposed by Hyland and Shaw (2016), namely *authenticity*, *groundedness*, *interdisciplinarity*, and *relevance. Authenticity* refers to the use of texts containing real examples of communication in the academic world, while *groundedness* is committed to linking research and pedagogy by drawing on insights from research to inform instructional practices. Next, *interdisciplinarity* characterises EAP as an application of an eclectic range of theories and methods to specific registers (i.e. text varieties associated with a particular situation of use) to provide an understanding of academic communication and classroom practice, and *relevance* manifests itself in EAP by relying on needs analysis to identify the specific areas including skills, texts or communicative practices that are relevant to a particular group of learners (Hyland & Shaw, 2016). These principles thus ought to enable EAP provision to facilitate success in effective written communication in academic contexts where British English is the desired variety. This is regarded as vital in academic settings since writing represents the main way of consolidating and demonstrating knowledge, thus performing a gate-keeping role with effective writing leading to success and ineffective writing to failure in the academy (Flowerdew, 2016; Hyland, 2013).

EAP thus seeks to identify and analyse the main genres frequently employed in academic contexts and clarify these for students through consciousness-raising and linguistic awareness tasks. This is intended to provide insights into the ways expert users of language achieve a variety of complex goals associated with their disciplines by focusing on the language use by members of various disciplinary cultures to manipulate generic conventions (Bhatia, 1997). Some argue, however, that due to this focus predominantly on expert texts, when these texts are used in academic writing classes, students often find them not only intimidating but also irrelevant to the texts they are required to produce (Wingate, 2018). Some researchers have, therefore, focused on genres produced in university settings to assist students in producing these types of texts (e.g., Bunton, 2002; Soler-Monreal et al., 2011; Thompson, 2016).

Thompson (2016), for instance, emphasises the value of a genre approach to PhD theses and MA dissertations which are often perceived as a challenge by students, owing to the length of the text they have to produce as well as the complexity of the rhetorical task. Specifically, students have to demonstrate not only their subject knowledge and an

ability to conduct independent research, but they also have to adhere to widely accepted conventions of communication within their disciplinary community. Thompson (2016) sees the benefit in employing a genre approach to this university genre in the ability to identify the communicative functions that thesis and dissertation writers need to express and in what sequences, followed by identification of the linguistic structures that typically perform the functions with consideration given to the specific community preferences. This is intended to assist student writers in producing this type of texts by providing them with a range of options that should enable them to select the option closest to their purpose. Such finding are thus valuable in EAP contexts as they offer insights into rather complex texts (Thompson, 2016). Several other researchers (e.g., Bunton, 2002, 2005; Soler-Monreal et al., 2011) have informed EAP practice by exploring the various sections of theses and dissertations for identification of typical moves and linguistic features that students can draw on during written production of these university writing genres.

In EAP settings, therefore, the goal is to demystify the genres relevant to students and help them understand and produce them effectively in writing. ESP genre analysis has assisted in this task by exploring the discourses of the various disciplines including their characteristic grammatical and lexical features. Such features have been shown to be prevalent in these contexts, which has led to recognising academic writing as a specific register within the ESP tradition. Some of the most common features of an academic register is the degree of formality achieved through the use of impersonal and nominalised style (2.2.3) and specialist vocabulary (Hyland, 2004), which is a specific type of vocabulary common across as well as within a variety of academic disciplines (3.2.3). ESP pedagogies in EAP contexts thus aim to raise students' awareness of the features generally associated with academic registers so as to help them become familiar with the genres relevant to their academic needs and subsequently enable them to produce these genres more effectively.

English for Academic Purposes at UK universities

EAP provision at UK universities encompasses pre-sessional, in-sessional and foundation courses (1.1) and targets primarily international students, whose first language is not English (i.e. L2 learners). This focus on overseas students suggests, however, that L2 writing is often seen as deficit and in need of remedial action. It also fails to consider that "[a]cademic language is … no one's mother tongue" (Bourdieu et al., 1994, p. 8),

meaning that students from all backgrounds entering tertiary education are novices in academic settings and often lack prior experience with and exposure to the type of writing they are required to produce (Tribble & Wingate, 2013).

Due to the focus of EAP provision on overseas students, EAP learners tend to be a heterogeneous group represented by diverse backgrounds, who bring to the academy their knowledge and experiences from their prior educational practices. It then follows that these different groups of students tend to use language in different ways in relation to, for instance, making persuasive appeals, ways of incorporating material, or drawing on different linguistic features; hence, the process of acquiring the norms of the kind of English prevalent in academic settings must also differ for various learner groups (Hyland & Shaw, 2016). The development of the skills and practices necessary in academic contexts is thus assisted by EAP provision.

2.2.3 Corpus linguistics

The corpus linguistics approach, often applied to genre investigation, has been most influenced by Biber (1988) and further developed by Biber and Conrad (2009), who approached text analysis from both a *genre* perspective and a *register* perspective. In their analytical framework, Biber and Conrad (2009) regard a *genre* approach as a way of characterising text varieties on the basis of the rhetorical structuring of texts, meaning that the linguistic analysis of texts from the genre perspective focuses primarily on the conventional structures used to construct complete texts. That is, the genre perspective is concerned with language characteristics that often occur only once in a particular place in a text but play a key role in the construction of texts. For instance, the genre organisation of an academic research article is characterised by Introduction-Methods-Results-Discussion (IMRD) sections marked by section headings which organise research articles by convention. Hence, genre features can be described as conventional rather than pervasive. For this reason, genre studies within this tradition are based on analysis of complete texts (Biber, 2019; Biber & Conrad, 2009; Egbert & Gray, 2019).

In contrast, a *register* approach is associated with a particular situation of use and is concerned with the functional relationships among the situational contexts, the communicative function and the linguistic forms of a text. Registers are described for their typical linguistic features comprising lexical and grammatical characteristics, based on the underlying assumption that core linguistic features are functional and are thus

used in association with the situational context and communicative purposes of texts. That is, a register analysis focuses on identification and description of linguistic characteristics which may occur in a variety of registers but are considerably more pervasive throughout texts in the target register. For instance, the use of questions and personal pronouns (e.g. 'I', 'you' and 'we') has been found to be among the linguistic features particularly common in face-to-face conversations compared to other registers (Biber & Conrad, 2009).

The genre and register perspectives thus differ in that complete texts are required for the description of the linguistic characteristics associated with the genre perspective focusing on the linguistic characteristics used to structure complete texts, while text excerpts are sufficient for a register analysis due to its focus on the frequent and pervasive linguistic features occurring throughout texts. It is noteworthy that complete texts analysed from the genre perspective can also be approached from the register perspective (Biber, 2019; Biber & Conrad, 2009; Egbert & Gray, 2019).

In corpus linguistics, register analyses of the various pervasive linguistic features in complete texts or text excerpts have often been conducted quantitatively using corpora, referring to large principled collections of texts intended to offer a representative picture of a genre or a group of writers under investigation (Hyland, 2016b). The methodologies of corpus linguistics enable the processing of these large numbers of texts (i.e. corpora) and their subsequent analyses for various textual features particularly frequent in texts representing a specific register (Biber et al., 1998).

Corpus linguistics is also one of the most commonly used contemporary approaches to the investigation of written academic discourse (Blagojević, 2016) and numerous recent studies have drawn on the methodology of corpus linguistics in the context of academic writing (e.g., Biber & Gray, 2016; Biber et al., 2011, 2013, 2016). These studies emphasise the importance of nouns in the academic register, describing the nominal style of written academic discourse. Biber and Gray (2016) present convincing evidence that the specific structures functioning as noun modifiers include attributive adjectives, nouns as nominal pre-modifiers, prepositional phrases as nominal post-modifiers, and appositive noun phrases. They also point out that "these features have been mostly overlooked in previous research on academic writing, despite their quite frequent use in this register" (Biber & Gray, 2016, p. 126). In terms of L2 academic written production,

Biber et al. (2011) found that L2 learners' progression manifests itself by moving to a denser use of noun phrase modifiers as they develop their academic writing skills.

The prevalence of the nominal style in the academic register is closely related to its heavy reliance on nominalisation, which is regarded as the most important manifestation of grammatical metaphor in modern science writing (Halliday, 1985, 2004; Halliday & Martin, 1993). Grammatical metaphor refers to the description of processes and actions, usually expressed by verbs, converted into nominalised forms expressed by nouns (Biber & Gray, 2016). Biber (1988) lists the following suffixes as characteristic of nominalisation: *-tion*, *-ment*, *-ness*, and *-ity* (including the plural form). The process of nominalisation also allows the addition of modifiers and qualifiers, thereby contributing to the concentration of information (Banks, 2005), in line with the above-noted use of complex noun phrases with phrasal modifiers characteristic of academic writing. Nominalisations are regarded as an important aspect of academic writing style due to the various functions that they have in academic texts, such as maintaining an impersonal tone or contributing to textual cohesion, which are seen as relevant in the production of quality academic texts (Baratta, 2010).

Numerous studies into various aspects of university student writing have also utilised the methodology of corpus linguistics by compiling learner corpora (i.e. collections of texts produced by learners) comprising language data produced by learners of various ages, linguistic backgrounds, proficiency levels or regions that can be used to explore a wide range of research topics (Díaz-Negrillo & Thompson, 2013). Staples and Reppen (2016), for instance, investigated a corpus comprising nearly 400,000 words of first-year university student writing across three L1s (English, Chinese, Arabic), two genres (rhetorical analysis and long argument) and language ratings with a particular focus on the vocabulary and grammar used by the student writers. Their findings show similarities in the use of lexico-grammatical features across writers from the different L1s, owing to their shared status as developing novice writers. Their study also revealed the impact of the genre on the lexico-grammatical choices of the novice writers. Similarly, various lexical and syntactic features in first-year university students' academic essays were explored in Hinkel's (2003) study drawing on a corpus of approximately 323,000 words representing six L1s (English, Chinese, Japanese, Korean, Indonesian, Arabic). Her investigation found that L2 students deploy a considerably higher number of simple syntactic and lexical features containing grammar and

vocabulary prevalent in spoken discourse than L1 students. Hinkel (2003) thus emphasises the importance of instruction for university-bound L2 students to concentrate on expanding their syntactic and lexical repertoire. The methodology of corpus linguistics was also utilised in Milton's (2001) study comparing features of writing produced by Hong Kong students of English during their three-year undergraduate programme with texts produced by UK students representing standard English on the basis of a corpus comprising approximately 1.5 million words of the Hong Kong university students' writing and a corpus of 510,000 of L1 students. One of the differences found by Milton was the different writing strategies used by the L1 and L2 students in his study. For instance, Milton's study reveals greater reliance of L2 students on the assignment prompt in the integration of lexical items in their writing with about 4% of all words in the Hong Kong students' corpus being formed by eight adjectives drawn from the assignment prompt, compared to only 0.5% in the UK students' texts. His study also showed that the least proficient Hong Kong students relied particularly heavily on the vocabulary items contained in the assignment prompts.

University student writing has also been investigated through the methodology of corpus linguistics at postgraduate and at doctoral level, generating insights into the deployment of multi-word sequences in various contexts. Hyland (2008), for instance, explored 4-word sequences (which Hyland refers to as 'academic clusters') in a corpus of 3.5 million words comprising master's theses, doctoral dissertations and research articles across four disciplines. He found that the three different genres display different cluster patterns resulting from the fact that master's students, doctoral students and professional academics draw on different resources in developing arguments in their writing.

Corpora can also be used on IFP programmes to explore various aspects of language use such as how words are used, the typical environments in which they can be found, or their frequencies of occurrence (Thompson, 2015a). One such study into foundation-level students' writing was conducted by Chuang and Nesi (2006), who explored a corpus of 88,000 words of academic essays produced by L1 speakers of Mandarin with a particular focus on identification of errors in the students' academic essays, which were subsequently categorised, quantified, described and explained. They found that the vast majority of errors were grammatical (approximately 86%), with the most common errors relating to the mismanagement of the article system (e.g. missing or redundant definite

or indefinite article or mis-selection between them). They speculate that the frequency of this type of error may be due to the students' L1, which does not have an article system, nor does it distinguish between countable and uncountable nouns or between singular and plural noun forms. Their study hence attempts to provide insights into the type of errors typically made by Chinese foundation students in their written English production with the aim of creating remedial grammar materials for Chinese students studying in the medium of English.

These studies demonstrate that corpus linguistics methodology has been utilised to investigate learner writing for various linguistic features at different levels of study, ranging from foundation to doctorate.

In sum, the corpus linguistics approach is characterised by its attempt to uncover typical patterns of use in a large collection of naturally occurring texts (i.e. corpora) prevalent in specific settings assisted by computer analyses (Biber et al., 1998). As corpus linguistics is not restricted to any particular aspect of language, one strength of this approach relates to the fact that it can be utilised with a wide range of applications in any area of linguistic research (McEnery et al., 2006), including the study of academic texts, where corpus linguistics has made invaluable contributions. However, a potential limitation of the corpus linguistic approach may be seen in the extent to which it allows writing researchers to generalise about populations as the findings are typically restricted to the sample forming the corpus under study. Therefore, a number of criteria need to be considered in the process of a corpus design, including size, balance (i.e. the range of text categories) and representativeness, which in turn impact the generalisability of the findings. Researchers thus need to consider the kinds of texts to include, the number of texts, the number of samples from each text, the selection of text samples from within texts, the length of text samples (i.e. the number of words in each sample), and the overall number of words in the corpus (Biber, 1993; Biber et al., 1998). A further potential limitation of employing the methodology of corpus linguistics may be seen in its focus on the writing product, neglecting the complex social factors surrounding writing practices.

2.2.4 Genre-based approaches combined

The three genre-based approaches representing the linguistic tradition in genre studies discussed in this chapter were shown to be socially recognised ways of using language with genre analysis resting on two central assumptions: a) the features of similar group of texts cannot be separated from the social context in which they are created and used; b) those features can be described in a way that relates a text to other texts like it as well as to both the choices and constraints placed on the producers of these texts. Some of their shared characteristics, therefore, include their view that linguistic characteristics of texts cannot be separated from social context and function. This is because genres are based in the societies that develop them; therefore, genres need to be mastered by individuals who seek to be accepted into the society and be able to engage with the community by being familiar with the rhetorical options available in the given circumstances together with knowing what is acceptable to say and in what order (Thompson, 2005). As has become apparent from the discussion above, the approaches differ, however, in their conceptualisation of *genre*. Despite these differences, nevertheless, they can be used in a complementary manner.

Hunston (2013), for instance, considered the relationship between SFL and corpus linguistics in a study of ideology in a popular science text. In her study, she argues that these two approaches may be viewed as parallel, divergent and complementary and are thus useful when combined in studying ideology. In particular, Hunston (2013) shows the usefulness of combining SFL concepts relating to grammar and corpus linguistics focus on lexis and phraseology, and notes that while SFL provides an explicit, multi-layered and detailed model to study texts in relation to their social context, corpus linguistics provides not only quantitative information for large numbers of texts, but also offers insights into the role that lexis and phraseologies play in the organisations of texts. The possibility of combining corpus linguistics and SFL is also demonstrated by Flowerdew (2003), who investigated the similarities and differences between novice and expert writing in the problem-solution rhetorical pattern commonly encountered in technical academic writing in a corpus comprising 250,000 words of both student and professional writing. Her study drew on corpus linguistics to identify frequently occurring key words and on SFL to categorise the lexical realisations in the texts. Flowerdew (2003) notes that combining corpus linguistics methods with SFL perspectives may be a useful way of exploring the generic patterns of problem-solution

texts as it enables the investigation of how lexico-grammatical choices are expressed by writers. A further example of combining SFL and the methodology of corpus linguistics includes genre-based studies of university student writing by Nesi and Gardner (2012), which led to a classification of assessed university student writing into genre families comprising individual finer-grained writing genres. Their taxonomy was based on the compilation and analysis of the British Academic Written English (BAWE) corpus (www.coventry.ac.uk/BAWE) which they developed in collaboration with other scholars as part of a four-year project investigating genres of student writing in higher education.

Genres in university student writing based on BAWE

Nesi and Gardner (e.g., Gardner & Nesi, 2013; Gardner et al., 2018; Nesi & Gardner, 2012, 2018; Nesi et al., 2008) applied a combination of genre-based approaches to the investigation of university student writing with the aim of specifying and classifying the various types of writing that students at British universities are required to produce for assessment purposes. In their work, Nesi and Gardner (2012) draw on the BAWE corpus comprising a collection of 2,858 texts collected from 1,039 students between 2005 - 2007 across four UK universities (Warwick University, Reading University, Oxford Brookes University and Coventry University), four levels of study (from first year undergraduate to taught postgraduate) and four broad disciplines (Arts and Humanities, Life Sciences, Physical Sciences and Social Sciences). The assignments forming the BAWE corpus were all awarded a minimum mark of 60%, corresponding to 2:1 in the UK system (based on the students' tutors' judgement), to ensure that they conformed to departmental expectations by meeting a certain proficiency standard. The BAWE corpus can thus be regarded as reflecting proficient university student writing represented by a wide range of written assignments typically produced at British universities.

Nesi and Gardner's (2012) analysis of the textual data in the form of written assignments forming the BAWE corpus was complemented by additional information such as the assignment titles and awarded grades. Other collected contextual data related to the participants and included their gender, year of birth, L1 and the number of years of British secondary education received. It is noteworthy that the students' linguistic background did not impact the decision about inclusion of their assignments as it was recognised that proficient writing in British universities does not emerge from a homogeneous linguistic background. Therefore, the BAWE contains writing produced by learners from various linguistic backgrounds representing approximately 40 L1s.

Further, the students' perceptions about the type of assignment they had submitted were explored through interviews together with interviews with staff focusing on disciplinary expectations. These interviews showed that the most common labels to describe the students' assignments used by both students and staff were either 'reports' or 'essays', often used interchangeably and often failing to acknowledge that different assignment types with the same descriptor likely require different organisational patterns as well as language uses (Alsop & Nesi, 2009; Nesi & Gardner, 2012, 2018; Nesi et al., 2005).

All the collected assignments were read several times with the aim of identifying similarities and differences in both the language use as well as the organisation of the assignments regardless of how the assignments were described by staff, students or the assignment briefs. In doing so, influenced by the SFL approach to genre in their investigation of the linguistic features associated with the stages of genres, Nesi and Gardner (2012, pp. 31-32) separated "decisions about genre (based on social purpose and staging) and cultural ways of doing things from decisions about register (based on the specific situation)" to classify the assignments. This was achieved by an analysis and grouping of the social purposes and stages of the almost 3,000 pieces of texts contained in the BAWE. This led to construction of 13 genre families comprising further finer-grained genres (Table 2.1).

Table 2.1: Writing genres derived from the BAWE corpus

Social purposes	Genre family	Examples of genres
Demonstrating knowledge and understanding	Exercise	calculations, data analysis, short answers
	Explanation	legislation overview, methodology explanation, environment report, account of a natural phenomenon
Developing powers of independent reasoning	Essay	challenge, commentary, consequential, discussion, exposition, factorial
	Critique	academic paper review, legislation evaluation, policy evaluation, project evaluation, review of a book / film
Building research skills	Literature Survey	annotated bibliography, literature review, review article
	Methodology Recount	data analysis report, field report, lab report
	Research Report	research article, student research project, dissertation
Preparing for professional practice	Case Study	company report, organisation analysis, patient report
	Design Specification	building design, product design, website design
	Problem Question	law problem, logistics simulation, business scenario
	Proposal	book proposal, business plan, research proposal
Writing for oneself and others	Empathy Writing	expert advice to industry, expert advice to lay person, information leaflet, job application, newspaper article
	Narrative Recount	accident report, account of literature search, biography

(adapted from: Gardner & Nesi, 2013; Nesi & Gardner, 2018)

The genre families (Table 2.1) were further explored with the assistance of corpus linguistic methodologies, namely multidimensional analysis and keyword analysis, which provided insights into the differences between genre families, between genres and across disciplines, primarily in terms of the lexicogrammar of the identified genres of university student assessed writing (Nesi & Gardner, 2012). It is noteworthy that although Nesi and Gardner's (2012) approach to writing genre classification was primarily influenced by the SFL approach to genre, their genre families differ from those typically produced in SFL contexts (2.2.1). This is because their aim was to group similar assignments rather than imposing a classification scheme developed for other contexts, namely writing across the curriculum in mainstream schools in the SFL tradition (Gardner & Nesi, 2013; Nesi & Gardner, 2012).

The genre families identified on the basis of the BAWE corpus are the most recent classification of assessed university student writing in the context of British universities. This categorisation of student writing offers insights into the similarities and differences between genres produced in various academic disciplines, at different levels of study as well as the characteristics of proficient university student writing, motivated by the recognition that very little was known about the writing university students produce (Nesi et al., 2005). Further, considering the size of the BAWE corpus, the range of data collected, and the thorough methodological procedures involved in the analysis and classification of the assignments, it can also be regarded as the most comprehensive genre classification to date in the context of assessed university student writing in the context of British universities. The BAWE corpus has thus made an invaluable contribution to both the teaching as well as research of university student writing by classifying the various genres or assessed academic writing required of students at British universities.

2.3 Academic literacies

Academic literacies is associated with the seminal work of Lea and Street (1998). Their approach is grounded in placing emphasis on academic literacy practices constituting reading and writing within disciplines, which academic literacies research regards as the central processes of learning in higher education. These literacy practices can be seen as social practices associated not only with different communities, but also with broader institutional discourses. This view of literacy practices thus considers the cultural and

contextual aspects of writing and reading practices, which can generate important implications for the ways student learning is understood by exploring the nature of academic writing practices in various disciplinary contexts (Lea & Street, 1998, 2006).

In their work, Lea and Street (1998) hence suggest that insights into the nature of academic learning can be gained by exploring the understanding that both academic staff and students have with regard to their own literacy practices, without any prior assumptions about the effectiveness or appropriacy of these practices. They consider this crucial in the development of a complex analysis of what it means to become academically literate as meanings are contested among all the parties involved, including the students, staff and institutions. An academic literacies approach to student writing, therefore, considers the wider institutional context of student writing from the perspective of social and cultural practices, as opposed to focusing merely on the deficits of student writing.

Primarily in response to the view of student writing as deficit with writing skills foregrounded by institutions only when "construed as a problem to be solved through additional or remedial support" (Lillis & Scott, 2007, p. 8), the academic literacies model influenced a shift in the way literacy practices in university contexts are perceived by moving away from 'problems' in academic writing seen as individual students' failure and recasting these difficulties as an institutional issue (Lillis & Tuck, 2016). Academic literacies thus aims to understand student writing by taking into consideration the complex nature of writing practices at universities with broader institutional discourses. Accordingly, the academic literacies approach has foregrounded several aspects of student academic writing previously ignored, such as the impact of power relations on student writing, the role of identity, and academic writing as ideologically-inscribed knowledge construction (Lea & Street, 1998, 2006).

Academic literacies hence considers student writing beyond their texts by encompassing individual participants' behaviours as well as the wider social and cultural meanings that the students bring to their literacy practices. This includes the investigation of academic norms and conventions and what counts as knowledge in any particular academic setting as well as institutional policy in relation to issues of power, authority and identity (Coffin & Donohue, 2012; Lea & Street, 2006; Lillis & Tuck, 2016). These insights can be gained through a dialogic approach to student written production by providing 'talkback' rather than 'feedback', which focuses on the students' text in process as

opposed to focusing on the students' text as a product (Lillis, 2003). This relates to the academic literacies principle in which "*practice* is privileged above *text*" (Lillis & Scott, 2007, p. 10), corresponding to the view of academic writing as a social practice.

This is also linked with the ideological stance towards the object of study in academic literacies, which has been referred to as 'transformative' rather than 'normative' (Lillis & Scott, 2007). The former focuses on the perspectives of the writers and the resources they bring to the academy, while the latter rests on a number of educational myths, such as the homogeneity of the student population, the stability of disciplines, or an emphasis on identifying academic conventions and exploring how students might become proficient (Kress, 2007). Such a transformative approach is regarded as crucial in understanding writers' practices as well as in opening up meaning- and knowledge-making spaces (Lillis, 2019).

This is also further emphasised by Wingate (2019), who highlights the need for a more wide-ranging transformation in order to provide adequate academic literacy support to all students, which would be achieved by the integration of academic literacy instruction into degree programmes, meaning that academic literacy practices would be focused on as part of subject lecturers' regular practices. A key aspect of this transformation must hence lie in academic literacy instruction being available to all students as opposed to targeting specific groups, based on the premise that all students new to university need to become familiar with the literacy practices of their specific academic disciplines (Wingate, 2019).

Due to the focus on literacy practices in specific contexts, academic literacies researchers put great emphasis on the student experience as they engage with genres and university meaning-making. This is enabled by drawing on ethnography as the primary empirical methodology of this approach, utilising a wide array of ethnographically-oriented data to inform research, ranging from textual data to interviews, discussions and observations of the practices involved in the production of texts and participants' perspectives on both the texts and practices. Despite drawing on various data enabling a comprehensive analysis, the ethnography-oriented methodologies employed by the academic literacies approach to the investigation of university student writing can, however, be also seen as a limitation of this approach. This is primarily due to the small number of participants typically investigated, which may not be considered representative of the various higher education contexts. Further,

since the focus of inquiry tends to be the text producer or meaning maker (Lillis & Tuck, 2016), academic literacies studies typically do not focus on the linguistic aspects of the genres that the students are required to produce, thereby backgrounding textual analysis and neglecting the importance of the linguistic features of university genres, which are unarguably an inseparable aspect of successful university student writing.

Numerous researchers have followed an academic literacies approach to the investigation of the practices surrounding academic writing in university contexts, illustrating the focus on the transformative stance as well as the reliance on ethnographically-oriented data collection methods. Tuck (2015), for example, investigated perspectives and experiences of disciplinary academic teachers with regard to their understanding of what was satisfactory, generative and meaningful in their practice around student writing in their specific contexts. Her study involved 14 academic teacher participants across disciplines in six diverse UK universities and was conducted by adopting an ethnographically-oriented approach utilising a range of data including semi-structured interviews with follow-up text-focused interviews, marked assignments complemented by other data such as assessment materials, audio-recordings and observed sessions. Tuck's (2015) study shows that what may be considered transformative for students in terms of academic writing cannot be separated from teachers' transformation. She thus emphasises the importance of nurturing the conditions for teacher transformation as well as of motivating students to engage with academic writing at a deeper level.

The academic literacies approach was also used to investigate the context of a writing circle in a university setting by utilising extracts from writers' texts complemented by discussions about the writers' concerns about their texts. The transformative aspect of the writing circle was that the participants were viewing both writers and writing as situated rather than judging it as good or bad (Chanock et al., 2015). The transformative approach encouraged by academic literacies can also be found in Kaufhold's (2015) study exploring the interaction of a master's student with their supervisor, particularly their negotiation of the standards required for thesis writing within their institutional context and their specific discipline. Her case study shows how, despite their lack of awareness of the academic literacies approach, the student and their supervisor applied certain aspects of the transformative approach, such as the negotiation of accepted institutional norms or the varied ways of knowledge-making. Similarly, evidence that

students considered the academic literacies approach beneficial, particularly in applying the principles to their peer-tutoring practice, was found in a study exploring how student-tutors (i.e. students supporting other students' writing) understand the academic literacies framework on a 'Teaching Writing module' offered to third-year Creative Writing students (Adams, 2015).

These studies highlight the central role of socially situated accounts of the contexts in which writing occurs and their importance for both pedagogy as well as research. In terms of pedagogical practice, the studies demonstrate the importance of involving students as "legitimate participants in curriculum development, in the organisation of teaching and learning activities, as well as in subject specific knowledge creation and research" (Hilsdon et al., 2019, p. 2). This can, hence, be seen as linking students, disciplinary practices and the procedures and policies of specific university contexts, thereby providing a space for a comprehensive response to university student writing encompassing student-, disciplinary-, and institutional- practices (Wrigglesworth, 2019). The studies also highlight the adoption of the transformative stance, which offers new perspectives on students' writing that have influenced writing practitioners in various contexts.

From a research perspective, the studies outlined above illustrate the various contexts in which student writing can be explored and exemplify the researchers' focus on literacy practices in context. The studies also demonstrate that this approach to the investigation of student writing is less concerned with the qualities of what tutors may regard as successful texts, and puts greater emphasis on the student experience as they engage with genres and university meaning-making. The reviewed studies also show that this is enabled by drawing on ethnography as the primary empirical methodology of this approach, utilising a wide range of ethnographically-oriented data to inform research.

In sum, the academic literacies approach to both research and pedagogy has made an invaluable contribution to the investigation of student writing by providing an important space for a critical exploration of what is involved in academic writing and what it means to be an academic writer in contemporary academia (Lillis, 2019). In doing so, academic literacies has contributed to the understanding of academic writing in several ways: it has provided insights into the gap in understandings between students and tutors with regard to academic writing conventions; it has highlighted the important role that

identities play in the student academic writing; and it has challenged the deficit approach to student writing. Hence, academic literacies is regarded as a key influence on EAP researchers, particularly in promoting the need to take account of writers' perspectives.

Despite the influence of academic literacies on EAP, however, the EAP approach is not entirely in line with the academic literacies approach. This divergence results from the fact that academic literacies considers the unitary notions of academic writing and of transferable academic writing skills as a "portable package of competencies" problematic due to the varying demands of not only individual disciplines, but often also individual practitioners within disciplines (Lillis & Tuck, 2016, p. 33). EAP providers are thus faced with having to find a balance between the academic literacies approach and addressing students' learning needs by helping them develop the literacy skills required in the various academic settings in which they need to operate. This interface between academic literacies and EAP has been found in both traditions emphasising the role of the language in the academy, in aiming to help students succeed as "writers and communicators in the increasingly globalised, English-dominant academy" by foregrounding the nature of academic conventions and putting emphasis on academic literacy as a situated practice (Lillis & Tuck, 2016, p. 36). Although academic literacies has found some common ground with the EAP tradition, it is noteworthy that the academic literacies' focus on writing as 'situated activity' (4.2.2) is in stark contrast to the corpus linguistic approach to the analysis of text (2.2.3), which focuses primarily on writing as 'completed activity' (4.2.1).

2.4 Chapter summary

This chapter has discussed some of the most common approaches to the investigation of student writing. These include three genre-based approaches, namely Systemic Functional Linguistics (2.2.1), English for Specific Purposes (2.2.2) and Corpus Linguistics (2.2.3), and how these approaches can be combined (2.2.4), including the most recent classification of genres of assessed university student writing in the BAWE project. Due to its influence on academic writing research, the academic literacies approach (2.3) has also been focused on. The discussion of the approaches to the study of university student writing has highlighted the most important characteristics of these approaches and the contributions they have made to the investigation of student writing. However,

it has also highlighted several potential weaknesses of these approaches in the context of university student writing research.

The discussion of the approaches to the study of university student writing has also led to the identification of a gap in the current body of literature. This omission has been found in the investigation of foundation-level students' assessed academic writing. With the exception of Chuang and Nesi's (2006) study (2.2.3.), no other studies have been identified that focus on any aspects of assessed academic writing produced by foundation-level students in the context of British universities. This study hence aims to address this important omission by investigating the assessed academic writing of foundation-level students in the context of a British university through a focus on academic vocabulary as the building blocks of academic texts (3.2.3). Specifically, this study will draw on the ESP approach (2.2.2) by focusing on this specialised type of vocabulary prevalent in EAP contexts represented by the International Foundation Programme (IFP) in this study (4.3.1). In addition, it will also draw on the methodology of corpus linguistics (2.2.3) by employing computer analyses to extract and further explore the most prevalent vocabulary items associated with the academic register. This study will also be assisted by the BAWE corpus (2.2.4), adopting the BAWE genre family classifications. Also motivated by the academic literacies approach (2.3), the present study will utilise participant interviews focusing on academic writers' perspectives. Considering the focus on academic vocabulary, the next chapter focuses on vocabulary research in general and academic vocabulary in particular.

Chapter 3: Vocabulary research

3.1 Introduction

Having previously established the vital role of academic vocabulary in university students' writing in this thesis (1.2), this chapter focuses on vocabulary research by providing a background to the various aspects relating to vocabulary, such as what counts as a word and the different types of vocabulary (3.2.1). This is followed by an overview of what is involved in knowing a word and how vocabulary is learned (3.2.2). The next sub-section is concerned with academic vocabulary as a specialised type of vocabulary, defining and outlining its crucial role in academic contexts (3.2.3). This leads to a discussion of academic word lists and their important role in vocabulary research (3.2.4) and an overview of research utilising these word lists to investigate learner writing (3.2.5). Next, an overview of vocabulary research relating to foundation-level students is provided (3.2.6). The chapter concludes with a summary of the main points covered, leading to the gap in the current body or research, which informs this study's research questions (3.3).

3.2 Vocabulary research

3.2.1 What counts as a word and types of vocabulary

In vocabulary research, one of the first decisions to be made is regarding what counts as a word. Orthographically, a word is generally agreed to be "any sequence of letters (and a limited number of other characteristics such as hyphen and apostrophe) bounded on either side by a space or a punctuation mark" (Carter, 1992, p. 20). In vocabulary research, however, a distinction is often made between a *type*, *token*, *lemma*, *word family* and *lexeme*.

A *type* refers to a unique word form in a text, meaning that each word form is counted once only regardless of the number of its occurrences in a text. A *token* (i.e. a running word) is defined as a single occurrence of a word form in a text and is counted each time it occurs in any given text (Brezina, 2018). A *lemma* consists of a headword and its inflected forms (i.e. plural, third person, singular present tense, past tense, past participle, *-ing*, comparative, superlative and possessive) and reduced forms (e.g. *n't*) belonging to the same part of speech (Kučera & Francis, 1967), while a *word family* contains a headword (i.e. a stem form or a base word) with all its inflected as well as

derived forms (with the derivates resulting from adding affixes, e.g. *-ness*, *-ly*, *-ism*) (Bauer & Nation, 1993).

To illustrate the difference between the word family and lemma principle, Gardner and Davies (2014) use the word family *proceed* subsuming the following members: *proceed* (verb), *proceeds* (verb or noun), *procedural* (adjective), *procedure* (noun), *procedures* (noun), *proceeded* (verb), *proceeding* (verb), *proceedings* (noun). According to the lemma principle, however, the following members would be counted separately: *proceedings* (a noun meaning 'records' or 'minutes'); *procedure* (a noun meaning 'technique') and its inflected plural form *procedures*; and *procedural* (an adjective meaning 'technical' or 'routine'). Next, a *lexeme* is defined as a "a group of related forms which share the same meaning and belong to the same word class (part of speech)" (Biber et al., 1999, p. 54); that is, lexemes include inflected forms of words but count homographs separately. For example, the polysemous noun *time* can carry different meanings (e.g. a period or frequency), which would be recognised by the lexeme approach by counting *time* as two lexemes.

Some vocabulary researchers have adopted further units for counting words by modifying some of the above units for counting; these include a *flemma* and *modified lexeme*. A *flemma* (i.e. family lemma) is an expanded version of lemmas, which unlike the traditional lemma does not separate parts of speech (McLean, 2018). For example, the adjective *developing* and the verb *developing* represent different lemmas owing to their different parts of speech but would belong to the same flemma as a result of their identical word forms. A *modified lexeme*, unlike the traditional lexeme which includes inflected forms of words but counts homographs separately, counts the headword in all its parts of speech and includes all inflected forms. For example, *list* includes the noun and verb *lists*, the verb and adjective *listed*, the noun and verb *listing*, and the noun *listings* (Browne, 2013).

Some of the above outlined ways of deciding what counts as a word broadly map onto Bauer and Nation's (1993) influential model of word family levels of affix criteria. According to their word family levels, Level 1 considers each word form a different word (i.e. a type). Level 2 encompasses words with the same base and inflected forms (i.e. lemmas). Level 3 counts the most frequent and regular derivational affixes, while Level 4 considers frequent, orthographically regular affixes. Level 5 adds regular but

infrequent affixes and Level 6 includes frequent but irregular affixes. Finally, Level 7 contains classical roots and affixes (Bauer & Nation, 1993).

It is also worth noting that a word is not always represented by a single lexical vocabulary item as multi-word units are sometimes counted together. These multi-word items are sequences of at least two words, variably referred to as 'lexical bundles', 'lexical phrases', 'formulas', 'routines', 'fixed expressions' or 'prefabricated patterns' with little agreement on their defining characteristics or what to call them (Biber et al., 2004).

In addition to considering different units of counting words, vocabulary researchers frequently make a further distinction between four kinds of vocabulary. These are *high-frequency* words which occur in various uses of the language and cover a large proportion of words in texts (i.e. approximately 80% of the running words in spoken or written texts); *technical* words which are closely related to a particular topic or subject area; *low-frequency* words which occur very infrequently and cover a small proportion of a text; and *academic* words which are common in different kinds of academic texts (Nation, 2001). The high-frequency words can be further extended for special purposes by systematically restricting them to specific topics or language uses; this type of vocabulary is referred to as *specialised vocabulary*. It is thus possible to have specialised vocabulary for speaking, letter writing or reading academic texts in a particular discipline, for instance. According to this view, both technical and academic vocabulary are sometimes also regarded as specialised vocabulary (Nation, 2001), which can be seen as the specialised vocabulary of English for academic purposes (EAP) (Coxhead & Nation, 2001).

Considering the focus of this study, academic vocabulary is further discussed in section 3.2.3. The next section focuses on the various aspects relating to vocabulary knowledge and learning, applicable to all types of vocabulary (i.e. high-frequency, low-frequency, technical and academic).

3.2.2 Vocabulary knowledge and learning

Receptive versus productive

A further aspect of vocabulary research concerns vocabulary knowledge, learning and usage with a frequently-made distinction regarding *receptive* versus *productive* vocabulary, also variously referred to as *active* versus *passive*, *comprehension* versus *production*, *understanding* versus *speaking*, or *recognitional* versus *actual* vocabulary

(Melka, 1997). *Receptive* vocabulary use refers to “perceiving the form of a word while listening or reading and retrieving its meaning”, whereas *productive* use involves “wanting to express a meaning through speaking or writing and retrieving and producing the appropriate spoken or written word form” (Nation, 2001, pp. 24–25). This distinction in vocabulary use resembles the receptive / productive types of knowledge closely related to the distinction often made between the receptive skills of listening and reading, and the productive skills of speaking and writing, where *receptive* involves the comprehension of the language input received from others through listening and reading, and *productive* carries the idea of producing language forms by speaking and writing in order to convey messages to others (Nation, 2001). Relating to receptive vocabulary knowledge is Krashen's (1982, 1989) ‘Input hypothesis’, according to which learners progress when they are exposed to ‘comprehensible input’ that is beyond their current linguistic level of competence.

Although the dichotomy between receptive and productive has been widely used in vocabulary research and pedagogy, some argue that this distinction is rather fuzzy. Melka (1997, p. 99), for instance, maintains that the “boundaries between the two notions are mobile and can move according to various linguistic or extra-linguistic factors", meaning that the distance between them is neither fixed nor permanent. Melka (1997, p. 99) thus proposes that the distance between receptive and productive ought to be interpreted as “degrees of knowledge and degrees of familiarity” and concludes that it is unclear whether the two notions should be seen as two separate systems dependent on one another or rather as one system used in two different ways, with the latter approach having certain advantages due to its regarding the two notions as highly interactive.

Knowing a word

Receptive and productive vocabulary use includes a variety of aspects relating to knowing a word. Nation (2001) proposes a model describing what is involved in knowing a word, which at the most general level involves form, meaning and use (Table 3.1).

Table 3.1: Aspects of knowing a word

Form/Meaning/Use		Receptive	Productive
Form	Spoken	What does the word sound like?	How is the word pronounced?
	Written	What does the word look like?	How is the word written and spelled?
	Word parts	What parts are recognisable in this word?	What word parts are needed to express the meaning?
Meaning	Form and meaning	What meaning does this word form signal?	What word form can be used to express this meaning?
	Concept and referents	What is included in the concept?	What items can the concept refer to?
	Associations	What other words does this make us think of?	What other words could we use instead of this one?
Use	Grammatical function	In what patterns does the word occur?	In what patterns must we use this word?
	Collocation	What words or types of words occur with this one?	What words or types of words must we use with this one?
	Constraints on use	Where, when, and how often would we expect to meet this word?	Where, when, and how often can we use this word?

(adapted from Nation, 2001, p. 27)

The above outlined aspects of knowing a word, as proposed by Nation (2001), are in line with Laufer's (1997) set of properties necessary to knowing a word, which include form, word structure, syntactic pattern, meaning, lexical relations, and collocations. Laufer (1997) notes that familiarity with all these features would imply knowing a word as is usually the case with an educated L1 speaker; however, knowing may be partial in the case of L2 learners as they may have mastered only some of the word's properties.

Vocabulary learning

The above outlined aspects of knowing a word can be learned in different ways. As in L1 vocabulary research, second language acquisition research often makes a distinction between *intentional* (i.e. explicit) and *incidental* (i.e. implicit) acquisition. Intentional or explicit vocabulary learning is a conscious and focused study of words and involves activities intended to result in vocabulary learning, while incidental or implicit learning leads to the learning of vocabulary items as a by-product of other activities (e.g. reading or listening), which are not specifically aimed at vocabulary learning (Ellis, 1999; Nation, 2001; Schmitt, 2000). Such implicit or incidental learning often involves inferring the word's meaning from context as a primary vocabulary skill, which can be achieved through extensive listening or reading (Nation & Waring, 1997).

Both extensive listening and reading provide exposure to large amounts of text for general understanding of content, often over an extended period of time, which the learners can understand relatively easily with a high level of comprehension as it is within their linguistic competence. Both these approaches have thus been found to

benefit learners' vocabulary acquisition resulting from a great deal of exposure to the most frequently occurring words in the texts (Nation & Waring, 1997). This is supported by Nagy (1997), who also emphasises the importance of exposure to large amounts of comprehensible input that would enable the learners to encounter the unfamiliar words repeatedly in context. He thus points out that a single encounter of a word in reading or listening will not lead to a considerable depth of word knowledge. Due to the need for large numbers of texts and repeated exposure to vocabulary items, Schmitt (2000) claims that incidental learning is slower, more gradual and lacking the focused attention of explicit teaching. Schmitt (2000) thus places emphasis on increasing the amount of exposure as one way to expedite incidental learning. Both extensive listening and reading have been researched in the context of second language acquisition, discussed next.

The impact of extensive listening on vocabulary gains was investigated by Chang (2012), for instance, whose study explored the effects of audio recordings of graded readers on Chinese university students of English over a 26-week period. Chang's (2012) study found extensive listening to be an effective way of enhancing not only L2 learners' vocabulary knowledge, but also their listening competence. The acquisition of L2 vocabulary through extensive listening in a university context was also studied by Vidal (2003), who focused on the effects of lecture listening on vocabulary gains in first-year university students. Vidal's (2003) study showed significant vocabulary gains from extensive listening to lectures as well as several predictors of vocabulary gains including learner proficiency, the frequency of the word occurrence, the type of elaboration accompanying the word, or the type of word (i.e. technical or academic).

The learning of vocabulary through extensive reading was investigated by Chang and Hu (2018), for example, who studied vocabulary learning rates of 62 Taiwanese senior high school learners immediately after reading ten graded readers and three months later. They found a correlation between the word frequency and the learners' retention rate, which was shown to be consistent in both lower and higher proficiency group of learners, whereby a higher word frequency corresponded to a higher retention rate. Similarly, in the context of Korean university students of English as a foreign language, Suk (2017) used graded readers to measure not only vocabulary acquisition, but also reading comprehension and reading rate over a period of 15 weeks and found a positive impact of extensive reading on all three areas. Also in the context of tertiary education,

the effects of extensive reading on 89 L2 Taiwanese students at a technological university were explored in relation to their reading attitude, reading comprehension and vocabulary size (Chen et al., 2013). The study found that extensive reading had a beneficial impact on the participants with regard to all three areas under investigation.

The above studies into extensive reading demonstrate its beneficial impact on vocabulary acquisition. Such incidental acquisition through extensive reading, compared to explicit teaching of vocabulary, can be regarded as more powerful as it is “unlikely that instruction accounts for anywhere near as much vocabulary growth as does incidental acquisition from context during reading" (Nagy, 1997, p. 75).

The benefits of exposure to reading texts for vocabulary gain have also been emphasised in academic contexts in relation to academic vocabulary acquisition, where numerous opportunities to read suitable texts are particularly important for the acquisition of academic vocabulary (Nagy & Townsend, 2012). This is supported by research showing that students who tended to read the least reported the most difficulty with academic words (Zimmerman, 1997). It is thus reasonable to assume that reading provides learners with opportunities to encounter academic vocabulary in contexts by providing information about the properties of a word and thereby contributing “to the learners’ knowledge about the multifaceted nature of words” (Zimmerman, 1997, p. 136). This is further corroborated by Krashen (2012), who believes that academic reading is more effective and efficient for academic vocabulary acquisition than instruction as reading a large number of academic texts exposes learners to the complex meanings and grammatical properties of academic words.

Related to the research into the impact of reading on vocabulary acquisition is research into source-based writing (also referred to as textual borrowing, reading into writing, writing from sources, reading-to-write constructs or integrated writing tasks) investigating the use of a source text during the writing process, particularly the integration of vocabulary items in written production. This was investigated by Plakans and Gebril (2012), for instance, who explored the use of sources and their function in the writing of 145 undergraduate students in a Middle Eastern university. Their participants used sources as a language repository as well as to generate ideas about the topic. In a later study, Gebril and Plakans (2016) also investigated the influence of textual sources on the lexical repertoire of 130 undergraduate students at a Middle Eastern university and found that textual borrowing (i.e. borrowing from source texts)

has a significant impact on lexical diversity. Their study thus calls for L2 writing instructors to train students on how to select vocabulary from source texts for subsequent integration in their writing.

Both the benefits and limitations of writing with and without a source text were investigated by Leki and Carson (1997), who found that the main disadvantages of writing without source texts, as reported by their participants, related to deficits such as lack of familiarity with the topic, lack of writing models or of vocabulary relating to the topic. Writing with a source text, on the other hand, was regarded as easier by the participants since a source text provided students with vocabulary, writing style and information, thereby "freeing the writer from the need to find appropriate words or to figure out the appropriate rhetorical form" (p. 56).

It is interesting to note that in addition to reading sources, some researchers have found students' reliance on other types of textual sources in integrating vocabulary in their writing. Flowerdew (2003), for instance, found that some of the vocabulary items investigated in her study of student writing could be traced back to the assignment guidelines. Similarly, Milton's (2001) study of UK (L1 speakers) and Hong Kong Chinese university students also showed that both groups of students relied on the vocabulary contained in the assignment prompts, but the Hong Kong students drew on the wording of the script to a much greater degree than the UK students. His study also showed that the least proficient Hong Kong students relied particularly heavily on the vocabulary items contained in the assignment prompts.

The above-outlined studies into extensive listening and reading as well as the closely-related source-based writing demonstrate the important role of incidental learning emphasising exposure to appropriate textual input in the acquisition and use of vocabulary. Nevertheless, Sökmen (1997) argues that such incidental encounters with words is only one way of facilitating vocabulary acquisition and provides several arguments for not focusing solely on implicit acquisition of second language vocabulary. These arguments include the slow rate at which learners can acquire vocabulary by guessing from context, the likelihood of errors and low comprehension of texts encountered primarily by lower-level learners with limited vocabulary knowledge, lack of consideration for individual learners' styles of acquiring new vocabulary and the fact that inferring from context does not necessarily lead to long-term retention (Sökmen, 1997).

Therefore, to address the potential limitations of learning vocabulary from incidental exposures, Sökmen (1997) suggests that implicit vocabulary learning should be accompanied by explicit vocabulary instruction. This is supported by Nation (2001), who also suggests that although learners may initially encounter new vocabulary through exposure to texts, this ought to be accompanied by conscious and intensive study. Specifically, Schmitt and McCarthy (1997) propose that the most frequently occurring words ought to be given explicit attention, while the wider range of less frequent vocabulary is likely to be best acquired through wide reading. They believe that explicit attention is a good way to acquire some initial information about the encountered word, but a great deal of exposure is necessary to become familiar with the more complex aspects of the word's register constraints or collocational knowledge, for instance, which is likely to be gained from reading (Schmitt & McCarthy, 1997). The benefits of explicit vocabulary learning are highlighted by Schmitt (2000, p. 120), who notes that "explicit learning focuses attention directly to the information to be learned, which gives the greatest chance for its acquisition" and concludes that for L2 learners, both explicit and incidental learning are necessary and thus ought to be regarded as complementary.

Vocabulary learning strategies

One way of facilitating vocabulary learning is by the use of different vocabulary learning strategies (Schmitt, 2000). These strategies have been classified using various taxonomies, reviewed by Schmitt (1997) who synthesised them under two categories: *discovery* vocabulary learning strategies for learning a word's meaning and *consolidation* vocabulary learning strategies for consolidating a word once it has been encountered. Among the discovery strategies are *determination* and *social strategies*, with the former referring to learners discovering a word's meaning by using reference material or guessing from context, for instance, while the latter involves asking someone who knows, such as a teacher. As far as the consolidation strategies are concerned, these also include *social strategies* such as group work, which is considered beneficial for several reasons: it promotes active processing of information; it enhances the participants' motivation through its social context; it helps prepare participants for team working outside the classroom; and it enables learners to use and manipulate language to a greater extent due to little instructor intervention. Further consolidation strategies include various *memory strategies* (e.g. the use of mnemonics), *cognitive strategies* including repetition and the use of mechanical means to learn vocabulary, or

metacognitive strategies used by students to control and evaluate their learning, which involves maximising the effectiveness of learning by, for example, seeking exposure to L2 input. Some of the resources that can facilitate this input through exposure include various reading sources or interaction with L1 speakers of English, for instance.

Schmitt (2000) suggests that rather than using isolated learning strategies, learners use a variety of vocabulary learning strategies concurrently and thus need to actively manage their strategy use. Accordingly, one of the current trends in teaching second language vocabulary points to encouraging independent learner strategies by helping learners to become aware of how they can continue to acquire vocabulary on their own (Sökmen, 1997).

Schmitt (1997) used his taxonomy of vocabulary learning strategies to explore vocabulary learning preferences of 600 Japanese learners of English. He found that the most-used discovery strategy was the use of a bilingual dictionary as a reference source and the most-used consolidation strategies included verbal and written repetition, reflecting the learning style encouraged by the Japanese school system. Learners' attitude towards a different type of a reference source, namely the thesaurus feature of Microsoft Word, for vocabulary acquisition was investigated in another study (Salehi & Habibi, 2015), which reported the following benefits of using a thesaurus: exposure to numerous different words with the same meaning; the ability to see and read these synonyms in a short period of time; making conscious or unconscious comparisons between the old and newly acquired words; and learning unfamiliar words. These studies hence show learners' positive attitude towards using reference sources as a discovery strategy for vocabulary acquisition.

Since the term 'vocabulary' encompasses various types of vocabulary (3.2.1), Schmitt's (1997) taxonomy of vocabulary learning strategies is also applicable to specific types of vocabulary (e.g. academic or technical vocabulary). One such study explored L2 university students' strategies for learning academic vocabulary (Huong, 2018). Findings from this study include asking teachers or friends as the most common social strategy, reading academic texts as the most frequently used meta-cognitive strategy, and a tendency to draw on various reference materials as a determination discovery strategy. Similarly, Lessard-Clouston (2008) explored L1 and L2 graduate theological students' strategies to learn the technical vocabulary of their discipline. Among the vocabulary

learning strategies reported by the participants were practising new words in the papers they were writing and consulting a theological dictionary.

Some of the social strategies of vocabulary learning have also been investigated in various settings. In the context of learning Spanish as a foreign language, Dobao (2014), for instance, compared vocabulary learning in pairs with learning in small groups and found that interaction in groups led to considerably more instances of vocabulary learning than pair interaction, mainly due to the opportunities of learners to draw on other group members' linguistic resources leading to more opportunities for vocabulary learning. More recently, Lin (2018) compared the effectiveness of group and individual work on Chinese university students' L2 vocabulary knowledge gain. Lin's study showed substantially higher vocabulary gains when working in groups compared to working individually. Specifically, some of the advantages of group work due to its interactive nature include receiving and giving information, discovering richer lexical information as well as richer support for retention.

Vocabulary learning processes

In addition to the above-outlined learning strategies which have been found to facilitate vocabulary learning, there are three main processes that may lead to vocabulary learning. These are *noticing*, *retrieval* and *generative use*. Noticing is an important first step in the acquisition of vocabulary and occurs when learners give attention to a vocabulary item as they become aware of its usefulness. This may be affected by a number of factors such as the salience of the item in a textual input. Noticing also occurs when learners deliberately study a word, look it up in a dictionary, have it explained to them by a teacher or guess from context, leading to the vocabulary item's comprehension. Once a new word has been noticed and comprehended, consolidation is strengthened by retrieval. This can be receptive or productive with the former referring to the perception of the form and retrieval of the word meaning when it is encountered in reading or listening, while the latter involves retrieval of the spoken or written form of the word in speaking or writing in order to communicate its meaning. This is followed by a generative stage, which occurs when previously met lexical items are subsequently encountered or used in ways which are different from the previous meeting with the item. This can also be both receptive and productive. Receptive generative use involves encountering a word used in new ways in reading or listening,

while productive generative use involves the production of the vocabulary in new ways and contexts (Nation, 2001).

As noted previously, the above discussed aspects of vocabulary research are applicable to different subsets of vocabulary, including academic vocabulary. Considering the context of this study, the next section discusses academic vocabulary in more detail. It first defines academic vocabulary, then focuses on the role of academic vocabulary in academic settings. This leads to the important role of academic word lists which are reviewed next, followed by an overview of studies which have utilised academic word lists to investigate academic vocabulary in learner writing.

3.2.3 Academic vocabulary

Academic vocabulary is often divided into vocabulary used primarily in a specific discipline, known as *technical terms* (Harmon et al., 2009), *technical vocabulary* (Fisher & Frey, 2008), *domain-specific* (Baumann & Graves, 2010), *content-specific* (Hiebert & Lubliner, 2008), *or discipline specific* (Townsend & Kiernan, 2015) academic vocabulary, and academic vocabulary commonly used across various disciplines, referred to as *general academic vocabulary* (Baumann & Graves, 2010; Townsend & Kiernan, 2015). General academic vocabulary is also variously referred to as *sub-technical vocabulary* (Baker, 1988; Cowan, 1974; Yang, 1986), *semi-technical vocabulary* (Farrell, 1990), *specialised non-technical lexis* (Cohen et al., 1979), *academic vocabulary* (Nation, 2001; Paquot, 2010) or *academic words* (Coxhead, 2000).

Cowan (1974, p. 291), for instance, refers to general academic vocabulary items as "context independent words which occur with high frequency across disciplines", similar to the definition proposed by Farrell (1990, p. 11): "Formal, context-independent words with a high frequency and / or wide range of occurrence across scientific disciplines, not usually found in basic general English courses; words with high frequency across scientific disciplines". Townsend and Kiernan (2015, p. 113) define these vocabulary items as "words that appear with much greater frequency in academic texts than in other types of texts, such as literary texts or popular media" and add that these words are "typically abstract, technical, nuanced, and / or densely packed with meaning". Similarly, Coxhead (2000, p. 218) refers to academic words as "lexical items [which] occur frequently and uniformly across a wide range of academic material", corresponding to Nation's (2001, p. 189) definition of academic vocabulary as "common

to a wide range of academic texts, and not so common in non-academic texts". These definitions of academic vocabulary are also in line with that of Baumann and Graves's (2010, p. 6) who define these vocabulary items as "words that appear in texts across several disciplines or academic domains". Paquot (2010, p. 28) is more specific in her definition of academic vocabulary in terms of its uses and suggests that for productive purposes academic vocabulary be defined as "a set of options to refer to those activities that characterize academic work, organize scientific discourse and build the rhetoric of academic texts".

From the definitions of academic vocabulary, it becomes clear that these vocabulary items are an indispensable aspect of academic texts. Nation (2001) points out that academic vocabulary accounts for a considerable number of words in academic texts, meaning that these vocabulary items represent high-frequency words in academic settings. Findings for the academic vocabulary content of academic texts vary between approximately 6% (Browne et al., 2013a), 8.5% (Xue & Nation, 1984), 10% (Coxhead, 2000) and nearly 14% (Gardner & Davies, 2014).

Considering the prevalence of academic words in academic texts, academic vocabulary can be regarded as important for both comprehension as well as production of these texts (Coxhead & Byrd, 2007). In terms of comprehension, academic vocabulary is widely recognised as a vital component of academic reading abilities (Corson, 1997) with insufficient knowledge of this type of vocabulary potentially compromising learners' ability to comprehend academic discourse (Donley & Reppen, 2001). In terms of production of academic texts, academic vocabulary is regarded as a key element of academic writing style (Hyland & Tse, 2007). Thus, academic vocabulary items are such words which students across disciplines encounter in their reading and should also be able to deploy in their written production.

This makes academic vocabulary a vital learning goal for learners of English of Academic Purposes (EAP) (Nation, 2001). As such, academic vocabulary is therefore considered to be a high priority and a useful learning goal for learners pursuing academic study in English (Coxhead & Nation, 2001), particularly in relation to high-stakes assessment, where "control of academic vocabulary, or the lack thereof, may be the single most important discriminator in the 'gate-keeping' tests of education" (Gardner & Davies, 2014, p. 305). Considering that written assignments are the main mode of assessment

in university settings (1.2), the productive usage of academic vocabulary in assessed writing can thus be regarded as a crucial component of academic achievement.

3.2.4 Academic word lists

Academic vocabulary research and pedagogy have been transformed by the creation of academic word lists containing the most frequently occurring vocabulary in academic contexts. Word lists creation is typically guided by one of the organising principles or units of counting, such as the word family principle or the lemma principle (3.2.1). In addition, the construction of academic word lists is further guided by one of two main views of academic vocabulary, reflecting the relationship between high-frequency and academic words (3.2.1). One view assumes that learners have already familiarised themselves with high-frequency words and regards academic vocabulary as falling outside these words. The other view is not based on the assumption that learners are already familiar with high-frequency words and encompasses all vocabulary items occurring with high frequencies in a wide range of academic texts (Dang et al., 2017).

In line with the distinction between general and discipline-specific academic vocabulary (3.2.3), word lists can also be divided into general academic word lists providing a compilation of academic vocabulary occurring across academic disciplines, and discipline-specific academic word lists containing academic vocabulary used in specific fields of study. Within these two categories exist compilations of academic vocabulary occurring in the written as well as spoken academic register (Therova, 2020a).

Among word lists relating to specific academic disciplines are several engineering word lists (Hsu, 2014; Mudraya, 2006; Todd, 2017; Ward, 2009), a number of medical word lists (Hsu, 2013; Lei & Liu, 2016; Wang et al., 2008), an Economics Academic Word List (EAWL) (O'Flynn, 2019), a technical words list for finance (Tongpoon-Patanasorn, 2018), an Environmental Academic Word List (EAWL) (Liu & Han, 2015), a Nursing Academic Word List (NAWL) (Yang, 2015), an Applied Linguistics Academic Word List (Khani & Tazik, 2013), a Chemistry Academic Word List (CAWL) (Valipouri & Nassaji, 2013), a Theological Word List (TWL) (Lessard-Clouston, 2010), a science-specific word list (Coxhead & Hirsch, 2007), or a Business Word Lists (BWL) (Hsu, 2011; Konstantakis, 2007).

In terms of the academic spoken register, the following word lists have been developed: the Hard Science Spoken Word List (HSWL) (Dang, 2018a) and the Soft Science Spoken

Word List (SSWL) (Dang, 2018b), the Academic Spoken Word List (ASWL) (Dang et al., 2017), the Vocabulary for Academic Lecture Listening word list (VALL) (Thompson, 2015b) and an English Spoken Academic Wordlist (Nesi, 2002).

In addition to these, academic word lists containing multi-word expressions have also been developed, including the Oxford Phrasal Academic Lexicon (OPAL) (McCarthy, 2019), the Academic Collocation List (ACL) (Ackermann & Chen, 2013), Phrasal Expression List (PHRASE List) (Martinez & Schmitt, 2012), and Academic Formulas List (AFL) (Simpson-Vlach & Ellis, 2010).

One of the first attempts at a compilation of the most commonly used academic expressions in writing across disciplines was the University Word List (UWL) (Xue & Nation, 1984), no longer considered relevant today as it was later updated and replaced by Coxhead's (2000) Academic Word List (AWL). Other compilations of general academic vocabulary include the Academic Keyword List (AKL) (Paquot, 2010), the New Academic Word List (NAWL) (Browne et al., 2013a), the New Academic Vocabulary List (AVL) (Gardner & Davies, 2014), and most recently the Oxford Phrasal Academic Lexicon (OPAL) (McCarthy, 2019). Consistent with the nature of the International Foundation Programme (IFP) under investigation in this study (i.e. generic as opposed to discipline-specific), these general academic word lists are discussed next. First, the AWL is focused on as the first compilation of general academic words since the UWL, leading to the review of the NAWL and AVL created as a response to the AWL, followed by the AKL and OPAL.

Academic Word List (AWL)

The AWL (Coxhead, 2000) is based on a corpus of approximately 3.5 million words drawn from 414 academic texts published between the early 1960s and the late 1990s representing science, arts, commerce, and law. Each of these four discipline areas contains approximately 875,000 words and is further sub-divided into 7 more specific subject areas. The creation of the AWL was guided by the following principles: frequency (i.e. members of a word family had to occur a minimum of 100 times in the corpus); range (i.e. the occurrence of a member of a word family at least 10 times in each of the four sub-corpora and in a minimum of 15 of the 28 subject areas); and specialised occurrence (i.e. exclusion of the 2,000 word families in West's (1953) General Service List of English Words).

The resulting AWL contains 570 word families organised into 10 sub-lists according to decreasing word family frequency. The AWL word families account for approximately 10% of the total words in the selected academic texts forming the source corpus. However, the AWL does not represent the four disciplines equally as its coverage of law (9.4%), arts (9.3%) and science (9.1%) sub-corpora is somewhat under-represented compared to the coverage of the commerce sub-corpus (12%). The 3% difference between the coverage of the commerce sub-corpus and the other three sub-corpora may have resulted from the presence of key vocabulary items which occur with high frequencies in commerce texts, such as *economic* or *finance* (Coxhead, 2000).

The above characteristics of the AWL and the methodology behind its compilation have been subject to critique. One of main criticisms is its foundation on the General Service List (GSL), which has been criticised for its age (resulting in the inclusion of words less frequent and relevant today such as *merchant* or *telegraph*, and the omission of words in frequent use today such as *computer* or *internet*), range, coverage, utility, organising principle around word families (e.g., Brezina & Gablasova, 2015; Engels, 1968; Richards, 1974), inconsistencies with words selected for inclusion, under-representation of words common in spoken English, and several methodological decisions involving subjective criteria including ease of learning, necessity, stylistic and emotional neutrality (Therova, 2020b).

A further flaw of the AWL has been pointed out by Hyland and Tse (2007), who found that the academic vocabulary items occur with different frequencies in different disciplines and behave differently across disciplines in terms of range, collocation and meaning; thus, the AWL is unlikely to be of equal value to students across disciplines. The AWL's word family principle has also been seen as its potential drawback as word families "often bring together forms with very diverse meanings" (Durrant, 2016, p. 51) due to the inclusion of derivationally-related forms. This can be illustrated by the AWL word family *react* (headword) subsuming the following members: *reaction*, *reactive*, *reacted*, *reactions*, *reactivate*, *reacts*, *reactionaries*, *reactivation*, reacting, *reactionary*, *reactor*, and *reactors*, demonstrating the differences in meanings between *react* (i.e. respond), *reactionary* (i.e. strongly opposed to social or political change), *reactivation* (i.e. to make something happen again), and *reactor* (i.e. a device or apparatus) (Gardner & Davies, 2014).

New Academic Word List (NAWL)

As a response to the AWL's limitations, Browne et al. (2013a) developed the New Academic Word List (NAWL). The NAWL contains 960 academic words drawn from an academic corpus containing 288 million words comprising the following three components: Cambridge English Corpus (CEC) Academic Corpus containing 248 million words (forming 86.3% of the source corpus) represented by academic journals, non-fiction, student essays and academic discourse; Textbooks corpus containing hundreds of top-selling academic textbooks totalling 36 million words (forming 12.6% of the source corpus); and Oral corpus containing 3 million words (forming 1.1% of the source corpus) from the Michigan Corpus of Academic Spoken English (MICASE) and the British Academic Spoken Corpus (BASE).

The NAWL was developed following one of the principles of Coxhead's AWL; that is, the exclusion of general high-frequency words. To eliminate general high-frequency words, the compilation of the NAWL was preceded by the creation of the New General Service List (NGSL) (Browne et al., 2013b), based on a modified lexeme approach, and containing approximately 2,800 high-frequency words drawn from a 273 million-word sub-section of the 2 billion-word Cambridge English Corpus (CEC). The NAWL together with the NGSL cover 92% of the academic corpus used for its creation, with the NGSL covering 86% and the NAWL covering 6% of the Academic Corpus. Considering that the NAWL was developed in conjunction with the NGLS, it is reasonable to assume that the NAWL is based on the same unit of counting as the NGSL; that is, a modified lexeme principle even though the compilers have not made this information explicitly available.

Some of the strengths of the NAWL, compared to the AWL, could be seen in the considerably larger size of the source corpus and the fact that it is more recent. One of its potential weaknesses, however, relates to its unit of counting based on modified lexemes, which collapses different grammatical classes, resulting in grouping words displaying different meanings. For example, the NAWL entry *articulate* carries a different meaning as an adjective and verb, or *integral* displaying a different meaning as a noun and adjective. A further major limitation of the NAWL lies in no precise information being made available regarding the methodological procedures involved in its creation, which hinders its comprehensive evaluation (Therova, 2020a).

New Academic Vocabulary List (AVL)

Another response to the AWL, in particular to the word family principle as a unit of counting, is the New Academic Vocabulary List (AVL) (Gardner & Davies, 2014), which was created using lemmas to eliminate the issue of different meanings resulting from the fact that grammatical parts of speech are not considered in word families (discussed above). The extraction of the AVL lemmas was based on a corpus containing 120 million words of academic texts from the 425-million-word Corpus of Contemporary American English (COCA) comprising the following nine academic disciplines: humanities; social sciences; history; education; law and political science; science and technology; medicine and health; business and finance; philosophy, psychology and religion (Gardner & Davies, 2014). Thus, in terms of size, range and representativeness, the AVL source corpus can be considered an advance on the AWL.

The following five considerations lie behind the compilation of the AVL: its basis on a representative corpus of academic texts covering a range of academic disciplines; the use of contemporary academic as well as non-academic texts for comparison purposes; a statistical derivation of the AVL from a large and balanced corpus of academic and non-academic texts; the reliability and validity of the AVL having been tested against academic and non-academic materials; and the use of lemmas, not word families (Gardner & Davies, 2014).

The creation of the AVL was then guided by dispersion, range, ratio, and discipline measure. According to the dispersion criterion (using the Juilland's D measure), the lemma had to reach a dispersion value of at least 0.8 (with dispersion values ranging from 0.01 indicating the occurrence of a word in an extremely small part of the corpus to 1.00 pointing to perfectly even dispersion in all parts of the corpus). The range principle was intended to ensure that the lemma occurs with at least 20% of the expected frequency in at least seven of the nine academic disciplines. The ratio criterion specified that the frequency of the lemma has to be at least 50% higher in the 120-million-word academic corpus than in the 305-million-word non-academic section of COCA in order to eliminate general high-frequency words from the AVL. The discipline measure was used to eliminate discipline-specific words by excluding words with an occurrence more than three times the expected frequency in any of the nine disciplines. In brief, the ratio criterion had the purpose of excluding general high-frequency words, while the other three criteria were used to eliminate discipline-specific words (Gardner

& Davies, 2014). These procedures resulted in a compilation of 3,015 lemmas forming the AVL.

Some of the strengths of the AVL thus include the size and detailed description of the source corpus together with the use of contemporary texts representing a wide range of disciplines, its basis on lemmas eliminating the issue of meanings caused by derivational relationships, and thorough methodological procedures. Despite these strengths, however, Durrant (2016) found that about half of the words in the AVL are used very little and items which are frequent vary across disciplines. Nonetheless, he also found that 427 out of the total of 3,015 AVL lemmas are frequent across 90% of disciplines, hence supporting the concept of general academic vocabulary common across disciplines, though at a smaller scale than the entire AVL. In addition, basing the AVL on COCA means that varieties other than American English are not represented.

Academic Keyword List (AKL)

The AKL (Paquot, 2010) was derived from a source corpus comprising professional as well as student writing totalling approximately 3 million words. The composition of the source corpus included professional academic writing containing around 2 million words drawn from the Micro-Concord Corpus Collection B (representing the domain of arts; belief and religion; science; applied science; and social science) and the Baby BNC Academic Corpus (representing humanities; politics, education and law; social science; science; and technology and engineering); and university student writing (around 1 million words) derived from the Louvain Corpus of Native Speaker Essays (LOCNESS) and the British Academic Written English (BAWE) Pilot Corpus representing language studies, social sciences, psychology and history (Paquot, 2010).

The methodological procedure involved in the identification of the AKL vocabulary items was based primarily on keyness to extract distinctive words; that is, words occurring with greater frequency in the source corpus compared to the reference corpus (constituting a corpus of fiction comprising nearly 2 million words) and subsequently by applying the criteria of range (only words appearing in all sections of the source corpus were considered) and evenness of distribution using Juilland's D value (which had to be higher than 0.8) (Paquot, 2010).

The resulting AKL contains 930 individual words and is a compilation of potential academic words which "has been named the Academic Keyword List (AKL) to emphasize

the fact that it is the output of a data-driven set of criteria, the first of which is keyness, and not a list of academic vocabulary in its functional sense" (Paquot, 2010, p. 55). As for the composition of the AKL, the largest group is formed by nouns (accounting for 38.17% of the AKL vocabulary items), which have been found to be particularly common in the academic register (Biber & Gray, 2016) and can be seen as closely related to the nominal style of writing regarded as an important aspect of the academic register due to its various functions in academic texts, such as maintaining an impersonal tone or contributing to textual cohesion (Baratta, 2010) (2.3.3). The next category is represented by verbs (25.05%), followed by adjectives (19.35%) and adverbs (9.35%) with the remaining 8.06% of the AKL items representing prepositions, conjunctions, pronouns, articles, determiners and ordinal numbers (Paquot, 2010).

However, as Paquot (2010) acknowledges, the AKL still needs to be validated as it is not a final product which would guarantee pedagogical relevance. Therefore, each AKL vocabulary item should be "subject to a careful corpus-based analysis to confirm its status as an academic word and establish how it is used in academic prose in terms of meaning, phraseology, and sentence position" (Paquot, 2010, p. 63). A further limitation of the AKL relates to the source corpus of student writing, which is skewed towards social sciences and humanities and may thus not be representative of academic vocabulary commonly used in the hard sciences. In addition, since the AKL was not based on a list of pre-existing general words, 57% of the AKL items are contained in the GSL. This inclusion of general high-frequency words thus indicates that the words in AKL are not unique to academic texts. However, Paquot (2010) argues that the high occurrence of the GSL vocabulary in the AKL emphasises the vital role that general vocabulary items play in academic registers. Another limitation is inherent in the keyness approach and relates to the selection of the reference corpus which may have contained potential academic words, resulting in an exclusion of these words from the AKL. Further, the size of the source corpus is relatively small compared to the AVL and NAWL.

Oxford Phrasal Academic Lexicon (OPAL)

The Oxford Phrasal Academic Lexicon (OPAL) (McCarthy, 2019) is a collection of four different word lists: written words, written phrases, spoken words, and spoken phrases. These four lists are intended to "provide an essential guide to the most important words to know in the field of English for Academic Purposes (EAP)" (Oxford Learner's Dictionaries, 2019). The most relevant to this study is the written words word list. This

list is derived from the 71-million-word Oxford Corpus of Academic English (OCAE) comprising academic texts published by Oxford University Press in four disciplines: arts and humanities, social sciences, physical sciences, and life sciences. Similar to the AVL and AKL, the OPAL is not based on a pre-existing list of general high-frequency words. The creation of the OPAL was based on a keyword analysis, which aimed to identify the most important words in academic contexts. This was intended not only to make OPAL "more than just a checklist of words that learners need to know", but also to assist learners in using appropriate language in their academic language production (Oxford Learner's Dictionaries, 2019). The written word list contains 1,200 individual words divided into twelve sub-lists (each containing 100 words and their part of speech) ordered from the most frequent words in academic writing.

The OPAL is, to my knowledge, the most recent compilation of academic expressions available. It is also conveniently linked to The Oxford Learner's Dictionary of Academic English, which is a learner's dictionary focusing solely on academic English, giving learners access to a powerful dictionary database. However, one of the potential weaknesses of the OPAL written word list concerns its collapsing of individual words with different parts of speech, resulting in individual entries displaying different meanings, as can be seen from the following examples: *current, relative* and *potential* each display a different meaning as an adjective and as a noun. In addition, similar to the NAWL, insufficient detail has been made available about the methodological procedures involved in its creation, such as the selection of the reference corpus used in the keyword analysis, hindering its more comprehensive evaluation (Therova, 2020a).

Academic word lists: summary

This overview of general academic word lists shows that despite their common aim of providing a compilation of academic vocabulary used across a variety of academic disciplines, significant differences exist between them. These relate mainly to their size, source corpora, methodologies, age and organising principles. Two distinct views of academic vocabulary (3.2.3) also become apparent; that is, a) seeing academic vocabulary as falling outside general high-frequency words with the assumption that learners are already familiar with these vocabulary items (adopted in the AWL and NAWL); b) not assuming that learners are already familiar with high-frequency words (adopted in the AVL, AKL and OPAL). A summary of the main features of the word lists is provided in Table 3.2.

Table 3.2: Overview of academic word lists

Word list	Source corpus & source corpus size (tokens)	Discipline representation	Methodology & criteria	Organisation, size & coverage
AWL (2000)	Source corpus size: 3,513,330 64% sources in New Zealand 20% sourced in Britain 13% sourced in the USA 2% sourced in Canada 1% sourced in Australia	4 disciplines: science; arts; commerce; law: each sub-divided into 7 subject areas	Specialised occurrence Range Frequency	570 word families 10 sub-lists covering 10% of the source corpus
NAWL (2013)	Source corpus size: 288,000,000 Academic Corpus of the Cambridge English Corpus: 248 million words (86.3%) Textbooks: 36 million words (12.6%) Oral corpus: MICASE and BASE: 3 million words (1.1%)	Not specified	Exclusion of NGSL items	960 words covering 6% of the source corpus
AVL (2014)	Source corpus size: >120,000,000 13,000 contemporary academic texts from the 425-million-word COCA published in the USA 85 million words from academic journals 31.5 million words from academically oriented magazines 7.5 million words from the finance sections of newspapers	9 disciplines: education; history; law & political science; humanities; social sciences; philosophy, religion, psychology; science & technology; medicine & health; business & finance	Ratio Range Dispersion Discipline measure	3,015 lemmas covering 13.8% of the source corpus
AKL (2010)	Source corpus size: 3,040,004 Professional writing corpora: Micro-Concord Corpus Collection B & Baby BNC Academic Corpus Student writing corpora: Louvain Corpus of Native Speaker Essays (LOCNESS) & British Academic Written Corpus (BAWE) Pilot Corpus containing British English	Professional writing: arts; belief & religion; social sciences; politics; science; history; technology & engineering humanities; education & law; science; applied Student writing: language studies; social sciences; psychology	Keyness Range Evenness of distribution	930 words covering 57% of GSL items and 40% of AWL items
OPAL (2019)	Source corpus size: 71,000,000 Academic texts published by Oxford University Press	4 disciplines: arts and humanities; social sciences; life sciences; physical sciences	Keyword analysis	1,200 words 12 sub-lists Coverage not reported

(adapted from Therova, 2020a)

Despite their wide-spread use in both vocabulary research as well as pedagogy, two major limitations have been levelled at academic word lists. First, several corpus-based studies investigating general academic vocabulary (e.g., Durrant, 2014, 2016; Hyland & Tse, 2007, 2009; Martínez et al., 2009) suggest that the label of 'general academic vocabulary' common across a wide range of academic disciplines may be misleading as "there may not be a usefully large set of vocabulary which is frequent across disciplines" (Durrant, 2016, p. 50). This argument subsumes two important points: a) general academic vocabulary items are not evenly distributed across disciplines; and b) general academic vocabulary items are used differently across disciplines (Durrant, 2014). This accords with Hyland and Tse's (2007) finding that words "often occur and behave in different ways across disciplines in terms of range, frequency, collocation, and meaning" (p. 235). Durrant (2016, p. 50) thus emphasises that "[w]ordlists are based on the premise that some words are likely to be more useful to learners than others", highlighting one of the main limitations of general academic word lists, which lie in the difficulty in meeting the lexical needs of students pursuing various disciplines.

Another limitation of word lists of general academic vocabulary is their reliance on single-item vocabulary, failing to reflect that collocational knowledge is one aspect of knowing a word (Table 3.1 in 3.2.2). This can result in displaying additional meanings in specific disciplines. For instance, Hyland and Tse (2007) found that *attribute* used as a noun meaning 'feature' is more common in sciences, while *attribute* as a verb meaning 'ascribe to' is more frequently used in engineering. Similarly, *volume* meaning a 'book' is considerably more common in social sciences than in engineering and science where *volume* is predominantly used to refer to quantity. Therefore, "[b]y breaking into single words items which may be better learnt as wholes, vocabulary lists simultaneously misrepresent discipline-specific meanings and mislead students" (Hyland & Tse, 2007, p. 247).

Nonetheless, despite these potential weaknesses of general academic word lists, they have made an invaluable contribution to applied linguistics research and pedagogy, and many researchers have utilised these word lists in their studies to investigate learner deployment of academic vocabulary in their writing (3.2.5).

The current study draws on the AVL, which has been selected due to its suitability for the purposes of the current study as well as its many strengths in comparison to the

other word lists presently available (further discussed in 4.2.1). Accordingly, 'academic vocabulary' in this study is defined as single academic vocabulary items frequently used in a variety of academic disciplines, including general high-frequency words commonly found in academic texts.

3.2.5 Academic vocabulary in learner writing

From the above-reviewed general academic word lists, the Academic Word List (AWL) and the new Academic Vocabulary List (AVL) have, to a varying extent, been used to investigate academic vocabulary usage in learner writing. This section begins with a chronological overview of studies utilising the AWL to explore university student writing, followed by longitudinal studies of university student writing employing the AWL. Next, studies drawing on the AWL in settings other than university contexts are focused on, then studies using the AVL are discussed. This section concludes with a summary of the reviewed studies, highlighting the gap in the current body of academic vocabulary research.

One of the studies utilising the AWL explored the relationship between the deployment of academic words and holistic scores of L2 learners' essays at an American university (Nadarajan, 2011). This study's findings showed no relationship between the usage of AWL items and holistic scores, but found that higher scores were awarded to essays containing a higher percentage of AWL words. In another study, the AWL was used to explore the effect of an academic context on the lexical choices of a group of English as an additional language students at a New Zealand university, and academic conventions employed by the participants to incorporate vocabulary into their writing (Coxhead, 2012). This investigation reports participants' awareness of the different tasks and vocabulary they needed for their academic study as well as of the impact of the academic audience on the participants' lexical choices. Coxhead's (2012) study also found that some learners tended to incorporate quotations in their writing, particularly when they lacked confidence in vocabulary knowledge, and also drew on dictionaries to check the meaning of the word before using it in quotations. Some participants reported using paraphrasing and summarising techniques as ways of incorporating words in their writing. This shows the students' reliance on reading and reference sources in the process of writing. Coxhead (2012) thus calls for academic support to provide advice on the readings and vocabulary that students ought to focus on as well as practice and modelling of

academic writing conventions including using quotes and paraphrasing techniques as ways of incorporating vocabulary into writing.

Academic vocabulary (represented by AWL items) from the students' perspective was the focus of Brun-Mercer and Zimmerman's (2015) study exploring advanced L2 learners' perceptions of academic vocabulary and their conscious decision-making processes relating to the integration of academic vocabulary in written production as well as the factors influencing these decisions, together with the techniques that helped the learners use these lexical items effectively. Their study found that although their participants were aware of the importance of academic vocabulary in their writing, they had difficulties deploying these vocabulary items effectively and appropriately as they were not always aware of the register of a word. They also found that their participants felt more confident in using new vocabulary after learning the vocabulary items in a meaningful way and after multiple exposures to the vocabulary in various contexts. Opportunities to use the newly learned academic vocabulary items productively and receiving feedback on the usage of the academic vocabulary items were also among the factors perceived by their participants as beneficial. In light of their findings, Brun-Mercer and Zimmerman (2015) call for explicit teaching of the register of newly met vocabulary, drawing students' attention to examples of academic vocabulary and its usage in appropriate texts and providing learners with ample opportunities to practise using the same vocabulary item in different contexts.

Among other studies using the AWL is Masrai and Milton's (2018) investigation into academic vocabulary size by devising a test to measure whether AWL knowledge is related to word frequency from general corpora as well as the contributions of academic and general vocabulary knowledge to academic achievement. Their findings show that a range of frequencies in general corpora are covered by the AWL; in other words, their results point to an intersection between the knowledge of the AWL items and knowledge of general English vocabulary. They also report a correlation between learners' academic performance and both general vocabulary size and academic word knowledge.

A number of studies used the AWL to explore university student writing from a longitudinal perspective. Among these is Storch and Tapper's (2009) study, which measured the number of academic vocabulary types and tokens at two different

times (week 1 and week 10) of an EAP course focusing on the writing of postgraduate students in Australia. Their study, which used a test / re-test design, found that by week 10 of the course, there was an increase in the usage of both AWL types and tokens. Storch and Tapper (2009) partially attribute the increase in the AWL usage to the students' exposure to academic texts due to the extensive reading required of postgraduate students. In addition, they speculate that the improvement in the AWL usage could also result from the delivery of the EAP course that the students attended, which placed great emphasis on academic vocabulary not only in the seminars but also in the teaching materials and feedback which the students received on their writing. Another study investigating changes in the deployment of AWL items in university student writing using a test / re-test design is Storch's (2009) study conducted over one semester (i.e. 12 weeks) at an Australian university. In contrast to Storch and Tapper's (2009) study, this study showed no change in the percentage of AWL items used in students' writing after 12 weeks. Storch (2009) speculates that the lack of improvement in the AWL vocabulary items over time may be due to the relatively short period of the study (i.e. 12 weeks) and the fact that the participants in the study were already rather advanced users of English, for whom improvement may be more difficult or take longer to achieve. Similarly, Xudong et al.'s (2010) explored changes in the usage of AWL items over a similar period of time in graduate students' writing at a university in Singapore, also using a test / re-test design. Their study showed a slight increase in the deployment of AWL items; however this increase was not found to be statistically significant.

Among longitudinal studies conducted over a longer period of time, also employing a test / re-test design, is Knoch et al.'s (2014) study conducted at an Australian university. Their study, spanning one year, calculated the percentage of AWL words as one of the measures of lexical complexity to assess the students' writing development. They found that while the students produced a greater range of vocabulary in general, there was no change in the percentage of the AWL items after one year. This study formed part of their larger study (Knoch et al., 2015), conducted over a period of a three-year degree study in the same Australian university, which also showed no significant changes in terms of the usage of AWL items over a three-year period. Knoch et al. (2014) speculate that the lack of improvement in the deployment of the AWL words over one academic year might be due to the nature of

the tasks used in the test for the study as the students did not have an opportunity to write about a discipline-related topic, which could have generated a higher usage of AWL items. In Knoch et al.'s (2014, 2015) studies, the participants also reported that they did not receive feedback on their written assignments, which may have contributed to the lack of awareness of areas for improvement and consequently a lack of improvement in their academic vocabulary. Knoch et al. (2015) also found that those participants in their study who did not think their writing improved attributed this primarily to insufficient writing practice.

Among academic vocabulary studies in contexts other than universities is Olinghouse and Wilson's (2013) exploration of academic vocabulary in three writing genres (i.e. narrative, persuasive, and informative writing) of fifth-grade learners aged 10-11 years by using the AWL to calculate the percentage of academic words per text. Their findings showed that the participants' usage of vocabulary was impacted by both the writing genre as well as the topic of their writing composition. They also found that the fifth-graders in their study deployed a very small number of academic vocabulary in their writing compositions regardless of the writing genre; that is, the compositions contained approximately 1% of academic vocabulary on average. Olinghouse and Wilson (2013) thus suggest that more explicit instruction is needed on both teaching academic vocabulary as well as on incorporating these vocabulary items into written production. Also in contexts other than university settings, Cons (2012) investigated the use and misuse of the AWL words in the writing produced by secondary-level English learners and redesignated fluent English-proficient learners in high and middle schools in Southern California. Her study finds a rare use of academic words by both groups. Cons (2012) emphasises a number of pedagogical implications arising from her study, including more exposure to academic vocabulary items through exposure to academic texts and frequent opportunities to read academic texts, the need for explicit instruction on effective integration of academic words in written production, and more writing practice in general and in the deployment of academic vocabulary in writing in particular.

Due to the later publication of the AVL (i.e. 14 years after the AWL), considerably fewer studies have utilised this academic word list in vocabulary research. Csomay and Prades (2018), for instance, used the AVL as a basis for exploring the relationship between the deployment of academic vocabulary and the perceived quality of

student writing as well as the differences in the usage of these words across text types, drafts and levels of instruction. They found not only significant relationships between the use of academic vocabulary and scores in some text types, but also differences in the way academic vocabulary is deployed across text types and a strong impact of rhetorical purposes on the amount of academic vocabulary used. Their further finding relates to the greater impact of the text type on academic vocabulary usage than that of a topic. Based on their findings, Csomay and Prades (2018) call for explicit teaching of academic vocabulary employing a contextualised and text-type specific or genre-based approach. Also drawing on the AVL, Durrant (2016) investigated the use of these academic lexical items in a corpus of university student writing with the aim of examining to what extent university students draw on the AVL across academic disciplines, levels and text types. Durrant's (2016) study reports some variations across text types with a more significant variation across disciplines. His study also shows that around half of the AVL items are used very little in student writing and frequently-deployed items differ across disciplines; however, he also found a small core of 427 AVL items which were frequent across 90% of disciplines.

From the studies utilising the AWL and AVL to gain insights into the deployment of academic vocabulary in learner writing, it becomes apparent that both academic word lists have been used to explore various aspects of student writing, such as the relationship between the deployment of academic words and holistic scores, the misuse of academic words, the impact of the audience on the integration of AWL items by students, the role of the AWL in different writing genres or its deployment to devise an academic vocabulary size test. In addition, several longitudinal studies have utilised the AWL to explore learners' improvement in the usage of academic vocabulary over time. All the reviewed longitudinal studies used a test / re-test design; in addition, they all used short pieces of writing (typically 300 - 500 words) completed under timed conditions and were conducted over varying periods of time from ten weeks to three years. The studies also yielded inconsistent results ranging from no increase in the usage of the AWL items over a relative short period of 12 weeks to considerably longer periods of one and three years, to an increase in the deployment of the AWL words over a period of 10 - 12 weeks. The reviewed studies using the AVL also focused on various aspects of student writing including the deployment of these academic vocabulary items across different text types, drafts

and levels. The focus of the above reviewed studies was primarily on university student writing with some studies focusing on lower educational levels (i.e. fifth-graders and secondary school learners). However, none of the identified studies investigated the usage of academic vocabulary in foundation-level students' writing. The next section, therefore, reviews existing research relating to vocabulary in the context of foundation-level students.

3.2.6 Vocabulary research and foundation-level students

A number of recent case studies relating to international foundation students' vocabulary usage have been reported by International Foundation Programme (IFP) practitioners. One such study investigated the challenges associated with the teaching and learning of key terms for history and politics (Hutton, 2016). In his study, Hutton (2016) outlines the use of Moodle's glossary feature to support the learning of discipline-specific vocabulary and argues that appropriate use of this activity can promote vocabulary acquisition. The facilitation of the acquisition of discipline-specific vocabulary on an IFP was also investigated by introducing students to a compilation of their own corpora of specialist terminology, which was intended to enable the students to discover the meanings that certain lexical items display in their own discipline (De Vries & Raffin, 2016). This study suggests that this model of learning is beneficial to students' learning as it promotes learner autonomy and intrinsic motivation by providing students with the tools they need to both interpret and produce language in their particular discipline.

Subject-specific vocabulary was also the focus of a study exploring the introduction of flashcards in relation to learners' retention of maths vocabulary, reporting that students enjoyed learning vocabulary in this manner and found it useful (Gurr, 2016). The teaching of key conceptual vocabularies with a particular focus on lexis associated with contemporary global issues (e.g. globalisation or sustainability) was investigated by Watson and Edwards (2016). The aim of their study was to help the students utilise a range of existing skills by enabling them to complete a variety of tasks independently outside the classroom and subsequently engaging the students in assessed group projects, in which they were required to demonstrate a range of academic skills such as extensive research or presentation skills (Watson & Edwards, 2016).

Subject-specificity in terms of vocabulary in the context of IFP was also investigated by Groves (2016), who argues that despite having numerous features in common with the wider field of EAP, foundation-level teaching differs from EAP teaching in two ways: first, foundation programmes focus on the delivery of other subjects alongside the EAP element; second, foundation-level students are novices in the academic community. This means that foundation students need to not only develop their academic skills in general, but they also have to familiarise themselves with discipline-specific concepts and content. Groves (2016) thus proposes equipping foundation students with the skills of identifying and recording the core vocabulary of their discipline so as to create their own academic lexicon, which ought to help them adapt to the needs of their particular discipline.

The above case studies in the context of IFPs highlight the trend of IFP practitioners to focus predominantly on discipline-specific vocabulary, reflecting the nature of these IFPs. This trend, however, does not consider the fact that some foundation-level courses are generic in their nature as opposed to being tailored towards a specific academic field of study. Moreover, these studies fail to take account of academic vocabulary in the context of high-stakes assessed academic writing. Hence, from the above overview of studies relating to vocabulary research in the context of IFPs, it becomes apparent that there is a lack of studies investigating academic vocabulary in the assessed academic writing produced by foundation-level students at UK-based higher education institutions.

3.3 Chapter summary and research questions

This chapter focused on various aspects of vocabulary including what counts as a word and the different types of vocabulary (3.2.1) as well as what is involved in knowing and learning a word (3.2.2). It also focused on academic vocabulary as a specialised type of vocabulary and its role in academic contexts (3.2.3). This was related to the importance of academic word list for academic vocabulary research (3.2.4) together with an overview of research utilising these word lists for the investigation of learner writing (3.2.5). This chapter also provided an overview of vocabulary research in the context of IFPs (3.2.6).

The different aspects of vocabulary research discussed in this chapter highlighted the gap in the current body of academic vocabulary research, which lies in a lack of

studies focusing on academic vocabulary in the context of foundation-level students. Despite numerous case studies focusing on the teaching of academic vocabulary on IFPs, primarily in subject-specific domains (3.2.6), no research has investigated productive general academic vocabulary deployed by international foundation students in their assessed academic writing or the students' perspectives on the learning and usage of these specialised vocabulary items.

This is an important omission which ought to be addressed due to the crucial role of academic vocabulary in university contexts, particularly in written production where this type of vocabulary is considered an indispensable aspect of academic writing style (3.2.3). This is closely related to the fact that written assignments constitute the main form of assessment at universities and insufficient knowledge of these vocabulary items is often associated with a gap in academic achievement (1.2). In addition, considering the important role of IFPs in the UK higher education settings (1.3), it is deemed important to gain insights into the deployment of academic vocabulary in the assessed academic writing produced by foundation-level students, and their perceptions of the factors contributing to their learning and subsequent productive use, so as to be able to assist them in the development of their academic writing as one of the academic literacy skills that foundation-level students need to develop if they are to succeed in their subsequent studies. The present study hence aims to address this omission in the current body or research by addressing the following research questions (RQs) (as outlined in 1.4):

1. What are the density and diversity of academic vocabulary in international foundation students' assessed academic writing across writing genres?

2. To what extent do the density and diversity of academic vocabulary in international foundation students' assessed academic writing develop over an academic year?

3. What are the students' perceptions of the contributing factors impacting the acquisition, deployment and development of academic vocabulary?

In the current study, 'academic vocabulary' is defined as single academic vocabulary items identified on the basis of the AVL.

The next chapter focuses on the methods employed in addressing these RQs.

Chapter 4: Methodology

4.1 Introduction

This chapter discusses the theoretical frameworks and approaches relating to the investigation of writing, which have informed the data collection procedures and subsequent analysis. It first outlines the different writing paradigms that provide writing researchers with options in regard to selecting appropriate methods of inquiry (4.2). It then focuses on how the data collection was conducted in line with the relevant writing paradigms (4.3), followed by a description of the analytical procedures involved in the data analysis (4.4). The chapter concludes by summarising the theoretical approaches and methodological procedures that informed the current study and their suitability for the investigation of the phenomenon under study (4.5).

4.2 Theoretical frameworks and approaches

Relating to writing research are particular views of writing, texts and writers, which shape the approaches to textual analyses. Hyland (2016b) notes that awareness of the various options available for studying writing and their relationship to key methodological designs may be helpful to writing researchers. Accordingly, he outlines six predominant writing paradigms, which inform approaches to studying writing. These are: *writing as cognitive activity* regarding writing as a thinking and problem-solving activity; *writing as expressive activity* seeing writing as a creative act of discovery and imagination; *writing as ideology* placing emphasis on power relations in social contexts reinforced in texts; *writing as social activity* focusing on discourse expressing community purposes rather than texts seen as objects; *writing as completed activity* describing the language in the finished static product rather than the process of writing; and *writing as situated activity* emphasising the physical and experiential contexts in which writing is produced.

The particular view of writing influences the methods selected for studying writing, which are "underpinned by philosophical assumptions about the nature of the world and how we can know it" (Hyland, 2016b, p. 121); that is, our ontological and epistemological assumptions that influence the specific methods used to collect and interpret data (Heaviside, 2017). Most relevant to this study is the view of *writing as*

completed activity (relevant to RQ1 and RQ2) discussed next and *writing as situated activity* (relevant to RQ3), discussed in 4.2.2.

4.2.1 Writing as completed activity

The view of *writing as completed activity* focuses on the writing product and is thus "limited to a static product not the result of a writer's dynamic effort to make meaning" (Hyland, 2016b, p. 122). This approach to writing utilises text analysis to explore various linguistic features in the finished product of writing with the aim of inferring rules of language usage. This has often been assisted by corpus linguistics methodology as this approach can offer a representative picture of a genre or a group of writers under investigation (2.2.3). This is achieved by enabling the processing of large principled collections of texts (i.e. corpora), which can be further analysed for various textual features (Biber et al., 1998).

In corpus linguistics, the distinction between *corpus-based* and *corpus-driven* approaches is often made (McEnery et al., 2006; Tognini-Bonelli, 2001), with the former referring to research starting with a pre-selected list of items to be examined for their frequencies and behaviour, and the latter being treated as the basis for any data-driven discovery (Baker & Egbert, 2016; McEnery et al., 2006). With corpus-based inquiry having become the norm in EAP research (2.2.3), corpus data can also be used in the context of IFP programmes to explore the various aspects of the student written production such as frequencies of occurrence or how words are used (Thompson, 2015a). Accordingly, a *corpus-based* approach to language investigation is adopted in this study as it will enable the identification and further investigation of pre-selected academic vocabulary items, which will form the basis of the textual analysis (RQ1 and RQ2). These academic vocabulary items will be identified using an existing list of general academic vocabulary (3.2.4).

From the reviewed general academic word lists (3.2.4), the Academic Word List (AWL) is not considered suitable for this study due to the issues inherent in its organising principle of word families, its basis on the General Service List (GSL), and its relatively small source corpus based on texts which are considered outdated (published between the early 1960s and the late 1990s). The New Academic Word List (NAWL) is also seen as unsuitable mainly due to its organising principle based on modified lexemes and a lack of information regarding the methodological principles and

procedures that guided its compilation. The Academic Keyword List's (AKL) weaknesses lie in its small source corpus size and the fact that it is not a final product which would guarantee pedagogical relevance. As for the Oxford Phrasal Academic List (OPAL), despite being the most recent compilation of academic vocabulary, insufficient information is available describing the methodological procedures involved in its creation.

Hence, the New Academic Vocabulary List (AVL) emerges as the most suitable word list for the purpose of this study. The rationale for this is multi-fold. First, the view of academic vocabulary not assuming that learners have already familiarised themselves with general high frequency words is relevant. This is because the participants are foundation-level students (4.3.3); in other words, the participating students are novice writers in English, meaning that the IFP is likely to be their first experience with academic writing in the English medium. Hence, it cannot be assumed that the learners have developed awareness of the function and the different meanings of general high-frequency words in the academic register. This is supported by Redpath (2015, p. 12), who notes that:

> "[w]hile students embarking on an IFP are likely to be familiar with many of the general high-frequency words from their prior English language learning experiences and may have a level of receptive knowledge of general academic vocabulary, their knowledge of the full range of general academic vocabulary is likely to require development".

Next, the size of the source corpus used for the compilation of the AVL is its strength in comparison to the majority of the other existing word lists. Further, its organisation on lemmas, as opposed to word families, will ensure a more accurate identification of academic vocabulary items in the student writing. Also, its representativeness of nine academic disciplines makes the AVL suitable for this study as it represents the wide range of academic domains to which the participants will progress on completion of the IFP (4.3.3). Hence, in this study, academic vocabulary is operationalised as individual academic vocabulary items identified on the basis of the AVL, which will form the basis of the textual analysis of the IFP students' assessed academic writing (RQ1 and RQ2) (4.3.4).

4.2.2 Writing as situated activity

Relevant to RQ3 is the view of *writing as situated activity* regarding writing as contextual performance impacted by the writer's personal attitudes and previous experience. In terms of ontological assumptions (concerned with the nature of reality) relating to this writing paradigm, the investigation of RQ3 draws on a relativist view of reality emphasising that every individual perceives reality differently according to their own subjective experiences of the world (Heaviside, 2017). This is reflected in this study by exploring the students' experiences with academic writing on the IFP and their perceptions of the main contributing factors that impacted the deployment and development of academic vocabulary items in their assessed academic writing.

In terms of epistemological assumptions (concerned with the nature of knowledge and the relationship between the researcher and the phenomenon under study), this study draws on constructivist epistemology associated with relativist ontological views holding that knowledge is constructed by the interactions that individuals encounter with the world. Hence, there are multiple knowledges as different individuals can assign different meanings to the same scenarios; that is, knowledge is constructed through people's meaning-making and is dependent on time and cultural location (Heaviside, 2017; Potter, 2006).

The above outlined relativist ontology and constructivist epistemology lead to the interpretivist research paradigm, which holds that people "actively interpret or make sense of their environment and of themselves, that the ways in which they do this are shaped by the particular cultures in which they live, and that these distinctive cultural orientations will govern what they do" (Hammersley, 2009, p. 19). Hence, interpretivists argue that it is not possible to gain an understanding of why people do what they do without gaining insights into their interpretations and meaning-making of their world, highlighting the emphasis placed on people's perceptions and beliefs (Hammersley, 2009). This emphasis on gaining understanding of people's perceptions is intended to provide insights into the IFP students' experiences with assessed academic writing and will be obtained through individual interviews and surveys (4.3.5).

4.3 Data Collection

This section describes the data collection procedures employed to address this study's research questions (1.4 & 3.3) in line with the writing paradigms (4.2). It first describes the research context (4.3.1) and the adopted recruitment procedures together with relevant ethical guidelines (4.3.2). It then provides an overview of the participants (4.3.3) and the collected data (4.3.4 & 4.3.5).

4.3.1 Research context

This study is set in the context of a generic (as opposed to discipline-specific) International Foundation Programme (IPF) at a UK university based in the South East of England and its two overseas campuses located in the United Arab Emirates (UAE) and Mauritius (MRU). The selected University is a public university offering a wide variety of undergraduate and postgraduate courses, which fall into one of the following categories: Art and Design; Business; Computer Sciences, Engineering and Maths; Construction and Architecture; Creative Writing and Media; Education Studies; Cinematics and Games; Healthcare and Social Work; Natural Sciences; Law; Performing Arts; Professional Practice and Work Based Learning; Psychology; Social Sciences; Sport and Exercise Science courses. In addition to these, the University offers discipline-specific foundation year courses preparing both local and overseas students for these degree programmes as well as the generic IFP in question. The IFP differs from the foundation year programmes in a number of ways: it is targeted at international students (referring to students from outside of the UK and EU[5]) who do not meet the academic requirements for direct entry onto the University's degree programmes; it accepts students with a lower English level than required for the discipline-specific foundation programmes[6]; it is not discipline-specific.

The goal of the IFP is to prepare these international students for undergraduate study at the University on a degree programme of their choice by helping them develop a range of skills necessary for degree study, such as effective communication skills, IT skills, research skills, critical thinking or the ability to work independently as well as

[5] At the time of data collection the UK was still part of the EU, meaning that EU students were not classified as 'international' students.

[6] The required language level for the IFP is IELTS 5.5 overall with a minimum of 5.0 in each of the test's four components (comprising reading, writing, speaking and listening), while the required language level for the discipline-specific foundation programmes is IELTS 6.0 or 6.5 overall with a minimum of 5.5 or 6.0 respectively in each of the test's components.

in teams. These skills are developed on four compulsory modules: *Academic Writing*, *Researching and Presenting*, *Developing Independent Learning*, and *Integrated Subject Based Project*. *Academic Writing* aims to develop the students' reading and writing skills; *Researching and Presenting* focuses on independent research techniques and oral presentation skills; *Developing Independent Learning* helps students develop strategies for independent learning and personal development planning as well as skills relating to the use of information and communication technology, mathematics and statistics for data analysis and presentation; *Integrated Subject Based Project* aims to enhance transferable skills development by using various concepts and techniques from a range of undergraduate disciplines.

These four modules are delivered over a period of one academic year (equalling 24 teaching weeks over six months), during which the students are required to attend 16 hours of taught sessions per week in the form of lectures, seminars and laboratories; typically, each module comprises a 1-hour lecture, 1.5-hour seminar and 1.5-hour laboratory. In addition, the students are expected to spend approximately 24 hours per week on individual study. The students' progress is assessed by summative individual and group assessments, both oral and written. The students are offered two assessment opportunities for each assessment task (i.e. in cases where the first assessment opportunity was unsuccessful, students are given one re-sit opportunity). The students have to pass all four modules, and thereby demonstrate that they have met the programme's learning outcomes. On successful completion of the IFP, the students can progress to a degree programme of their choice within the University.

4.3.2 Recruitment procedures and ethical considerations

For participant recruitment purposes in this study, the 'Ethical Guidelines for Educational Research' published by the British Education Research Association (BERA) (2018) and the 'Recommendations on Good Practice in Applied Linguistics' as outlined by the British Association for Applied Linguistics (BAAL) (2016) were followed. The specific guidelines related primarily to responsibilities to participants and stakeholders. Accordingly, relevant information was provided to participants and stakeholders in order to obtain their consent. Specifically, prior to approaching potential participants, ethics approval was first sought from The Open University Human Research Ethics Committee (Appendix 1), followed by obtaining ethics

approval from the target University[7]. Next, approvals from the UK-based IFP programme leader and the programme co-ordinators at both overseas campuses were sought[8]. Then, all potential participants were approached directly by the researcher either via email or, in case of some of the UK-based participants, in person. All students were provided with a Consent Form (Appendix 2) and an Information Sheet (Appendix 3) detailing the purpose of this study and other relevant information, such as assurance of confidentiality and anonymity or withdrawal procedure.

The sample for this study is formed by IFP students at the UK-based campus (2014 - 2017 cohorts) and the two overseas campuses (2017 - 2018 cohorts)[9], further described in 4.3.3. To ensure anonymity, each recruited participant was assigned a code identifying their location (i.e. UK, UAE, MRU), the year of the cohort and a letter to differentiate individual students. For example, MRU17/18/A indicates a Mauritius-based student from the 2017 - 2018 cohort who was identified with a letter A. Despite the different environments providing different opportunities for exposure and language use outside the classroom, the inclusion of IFP students from different cohorts and campuses is based on the rationale that the structure of the IFP was consistent across the three campuses over the period of data collection, meaning that the different IFP cohorts forming the sample in this study followed the same curriculum as well as assessment procedures.

Since the sample is formed by participants from different cohorts, different Information Sheets were distributed. For IFP students pursuing the IFP at the time of data collection (i.e. the 2016 - 2018 cohorts), the Information Sheet contained information relating to their voluntary participation in an online survey (non-UK based students only) and an interview relating to their experiences with academic writing on the IFP. This information was excluded for former IFP students based on the assumption that they might be less able to accurately recall their past experiences given the time that had passed since their completion of the IFP. The recruited participants forming the sample for this study are further described next.

[7] The Ethics Approval from the target University is not included in the Appendix so as to preserve anonymity of the University.

[8] As in footnote 7, this information is not provided in the Appendix to preserve anonymity of the University.

[9] The UK-based IFP had one cohort of students per academic year (October - April), whereas the two overseas campuses had two cohorts per academic year (October - April and January - July).

4.3.3 Participants

The recruitment procedure described above led to the recruitment of 193 participants across the three campuses. However, the representation of the different campuses is not symmetrical; the UAE-based students form the biggest proportion of the sample (113 students, i.e. 58.5%), the Mauritius-based students account for 55 out of the 193 participants (i.e. 28.5%) and the UK-based students represent the smallest proportion (25 students, i.e. 13%) of the sample. It is noteworthy that the vast majority of the participants were not permanent residents in their place of study. That is, the students typically arrived in order to attend the IFP during term-time (i.e. 24 teaching weeks over 6 months) and left after its completion. The exception were the Mauritius-based students, some of whom (21 out of 55 students) were permanent residents in their place of study. Despite the different learning environments, the composition of the sample is not believed to have had any effect on the textual data analysis or findings as all three campuses followed the same curriculum and assessment procedures, noted above.

The recruited participants were aged 16-26 (M=19, SD=1.71) and the majority intended to study on Business, Law, Computer and Psychology courses on successful completion of the IFP. The participants came from various linguistic, educational and ethnical backgrounds representing approximately 54 nationalities and 55 language backgrounds (L1s). It is noteworthy that some students stated more than one L1 and / or nationality and some failed to state either; hence, the figures relating to the students' nationalities and L1s indicate the minimum number of known L1s and nationalities. An overview of the participants is provided in Appendix 4 showing the student representation of the various cohorts across the three campuses, the students' age, gender and intended programme of undergraduate study together with an overview of the students' nationalities and L1s.

The different linguistic backgrounds represented in this study lend themselves to the investigation of intergroup homogeneity, referring to "similarities among L2 writers regardless of language background" (Crossley & McNamara, 2011, p. 271), with L2 referring to users of English who do not consider English their L1. Such an approach is important as it:

> "affords an examination of linguistic features inclusive to L2 writers as a composite …[and] can provide evidence that some aspects of L2 writing may not be cultural or independent, but rather based on the amount and type of linguistic knowledge available to L2 learners as a result of language experience and learner level" (Crossley & McNamara, 2011, p. 272).

It is noteworthy that a number of participants identified English as their L1 (4 UK-based students, 24 UAE-based students, and 13 MRU-based students) thereby not conforming to the definition of intergroup homogeneity relating to L2 speakers. These students are also included in the sample as the majority of the participants with English as their L1 come from places where non-dominant varieties of English are spoken (e.g. the Philippines, India, Mauritius, Nigeria, Jamaica, South Africa, Sri Lanka, Tanzania or Zambia) as opposed to the US or UK variety. Some of them have also stated another L1 in addition to English (e.g. Creole, French, Hindi or Hausa). Thus, considering the participants' various linguistic backgrounds, the sample is seen as suitable for investigation of intergroup homogeneity. This investigation will be based on textual, interview and survey data, discussed next.

4.3.4 Textual data

Selection criteria

In line with the view of *writing as completed activity* (4.2.1), the collected textual data were summative written assignments (submitted to the University for assessment purposes) representing the finished writing product that is the focus of investigation in this writing paradigm; hence, drafts were not considered for the purpose of this study. In addition, the textual data collection was guided by the following criteria: the written assignments were completed by individual students (as opposed to group projects which do not give all learners equal opportunities for written production), and they constituted continuous prose (as opposed to other types of assignments that the students were required to complete, outlined in Table 4.1 below). These criteria were used to eliminate assignments which would not be relevant to the aims of this study, such as oral presentations. Further, only first submissions were collected regardless of the awarded mark, including assignments awarded a fail mark due to not meeting the marking criteria sufficiently. This was to ensure that the same number of assignments was collected from all participants, in order to reflect the

students' linguistic competencies at approximately the same point of time in the academic year. An exception to this were submissions that individual markers had considered to be plagiarised initially on the basis of percentage of similarities generated by Turnitin (www.turnitin.com), which served as the University's assignment repository and plagiarism detection system. Assignments with high levels of similarities were further investigated for the extent of plagiarism by the markers. Submissions deemed plagiarised were not marked, nor was feedback provided on any aspects of these assignments. These submissions were excluded from this study on the premise that extensively copied sections of texts do not reflect the students' language production.

Since no feedback was provided on any aspect of plagiarised assignments, re-submissions of these assignments (typically submitted within a couple of weeks after the first submission) were also included in the sample since they can be regarded as the students' first submission which reflects their own writing. An overview of all summative assignments (following the University's assignment classification) is provided in Table 4.1, with the assignments meeting the above outlined selection criteria highlighted in bold.

Table 4.1: Overview of summative assignments

Assignments	Assignment type	Module	Submission October cohort	Submission January cohort
Assignment 1	Individual oral presentation	Researching and Presenting	November	March
Assignment 2	**Individual essay**	Academic Writing	November December	March April
Assignment 3	Individual information leaflet	Developing Independent Learning	November December	March April
Assignment 4	**Individual report**	Integrated Subject Based Project	January February	May
Assignment 5	**Individual report**	Developing Independent Learning	February March	May June
Assignment 6	**Individual essay**	Academic Writing	March	June
Assignment 7	Group report	Researching and presenting	April	June July
Assignment 8	Group presentation	Integrated Subject Based Project	April	July
Assignment 9	Learning diary & Listening task	Developing Independent Learning	April	July

From Table 4.1, four assignments (classified as essays and reports by the University) meet the above described selection criteria. These assignments were collected in an electronic format from Turnitin and were assigned the same code as individual

students (4.3.2). It is noteworthy that the assignments were collected after the academic year, meaning that this study was conducted independently from the teaching during the IFP; thus, neither the assignment instructions nor the actual teaching were influenced by this study. The collected assignments form a compilation of 772 texts (i.e. four assignments from each of the 193 participants) totalling 1,067,347 words (tokens) representing a range of topics (Table 4.2).

Table 4.2: Overview of collected assignments

Location **Cohort**	**UK** **2014-15**	**UK** **2015-16**	**UK** **2016-17**	**MRU** **2017-18**	**MRU** **2018**	**UAE** **2017-18, 2018**
No. of texts	**x12**	**x9**	**x4**	**x26**	**x29**	**x113**
Assignment 1 (essay)	Learning styles	Multiple intelligences	Learning styles		Multiple intelligences	Fake news
Assignment 2 (report)	Business organisational structures		Sustainable tourism			Sustainable initiatives; Business organisational structure; Ethics in social psychology
Assignment 3 (report)	Managing barriers to sustainability					
Assignment 4 (essay)	Social media and crime	Social media for academic purposes	Social media and crime		Social media for academic purposes	Surveillance society

It has to be acknowledged, however, that many instructors tend to refer to different types of writing "as either 'reports' or 'essays', often interchangeably and without acknowledging that different assignments with the same descriptor might require different uses of language and different organisational patterns" (Nesi & Gardner, 2018, p. 52). Thus, the collected assignments were further examined to provide a more accurate classification into writing genres in accordance with Nesi and Gardner's (2012) categorisation of university student writing (2.2.4).

Writing genre classification in this study

Nesi and Gardner's (2012) classification scheme has been adopted in this study as it offers a comprehensive categorisation of university writing genres based on relatively recent British university student assessed academic writing (2.2.4). In accordance with their classification scheme comprising 13 genre families, a close reading of the assignment briefs and the students' assignments led to the categorisation of the students' texts into three genre families: Essay (comprising finer-grained writing genres of Exposition and Discussion Essays), Problem Question, and Research Report.

In this study, an assignment brief refers to "the requirements of the assignment" (Gustafson-Pearce, 2009, p. 2585) as outlined in the assessment handbook with which the students were provided. That is, the brief "tells students what to do, involving skills, content, and thinking" (Dougherty, 2012, p. 23). The difference between the assignment brief and rubric is important to note, with the latter referring to assessment criteria or "scoring guides" (Dougherty, 2012, p. 24) providing a list of criteria and gradations of quality with varying levels from excellent to poor (Andrade, 2000, p. 13).

All titles under Assignment 1 (Table 4.3) were classified as 'Exposition Essays' as they shared the structure of 'thesis - evidence - thesis restatement' characteristic of this writing genre (Nesi & Gardner, 2012). For instance, in Assignment 1 on Learning Styles, the students analysed their learning style and provided evidence supporting their argument relating to the selected learning style. Similarly, Assignment 1 on Multiple Intelligences required students to analyse their intelligence type and provide evidence supporting their argument in relation to their type of intelligence. Assignment 1 on Fake News, although different from the other two in that it did not relate to the students' past experience, also required evidence supporting the stated position.

Table 4.3: Assignment 1 genre classification

Assignment topic/title	Assignment brief extracts	Writing genre features identified in student texts and brief	Assigned writing genre
Learning Styles	What kind of learner am I? Outline the term 'learning style'. Analyse your own learning style and how it relates to the characteristics of the learner types identified in Kolb's theory. Your points should be illustrated using examples from your past learning experiences. Illustrate points with experiences from past learning.	Three stages: thesis – evidence – restatement of thesis Statement of thesis in the introduction Series of arguments providing evidence to support the thesis; evidence argues for a stated position from the outset Conclusion summary restating the thesis in the context of the supporting arguments	Exposition Essay
Multiple Intelligences	Explain the theory of multiple intelligences and how this relates to learning. Analyse which of Gardner's seven intelligences reflect your own abilities best. Your answer should include examples of your learning practice illustrating how your approach to learning fits with Gardner's theory. Illustrate points with experiences from past learning.		
Fake News	Is fake news harmful? Your essay should: • Discuss arguments either for or against • Argue ONE side of the motion only • Directly answer the question, drawing a firm conclusion from your findings		

Assignment 2 titles (Table 4.4) corresponded to the writing genre of 'Problem Question' due to their aim to "provide practice in applying the theory and methods of the discipline" (Nesi & Gardner, 2012, p. 175) with a typical 'situation / context - problem - solution – evaluation' structure. However, a close reading of the students' texts belonging to this writing genre revealed some differences in their structure and organisation (e.g. the use of different headings). This likely resulted from the content reflecting the various titles, which represented topics common in specific fields of study (e.g. business or psychology).

Table 4.4: Assignment 2 genre classification

Assignment topic/title	Assignment brief extracts	Writing genre features identified in student texts and brief	Assigned writing genre
Business Organisational Structures	Create an organisational structure for a company that will start and deliver a project that will benefit the environment in your country. Research two companies that have environmentally beneficial projects as part of their operations and compare the organisational structures of both companies. Use this research into real world companies to inform and justify the organisational structure of the company that you are planning. Justify your structure design in terms of business theory explored in [the module] and your own secondary research.	Stages: situation / context - problem - solution - evaluation Although the given problem / scenario may be based on a real professional case, the task is designed solely to provide practice in applying theory and methods to the discipline To enable the students to concentrate on a single issue The main focus is on the process or arriving at the conclusion rather than the conclusion itself Converge on the best answer to a problem or dispute Knowledge demonstration without having to say anything original	Problem Question
Ethics in Social Psychology	Research one psychology case where ethics was an issue and write about the reasons these issues create a problem within social psychology. You will use this research into a past case, along with the social research ethical guidelines studied in class, to inform and justify your suggestions for resolutions to the ethics issues outlined in the case you have chosen.		
Sustainable Tourism	What is tourism and how does it affect people and the economy of a country? Why is sustainable tourism important? Choose an organization within the tourism industry. Discuss ways in which the organization demonstrates commitment to sustainable tourism. What are the main barriers in your chosen organization that prevents them from being more sustainable? Research, discuss and suggest implementable ideas to enable the organization to support sustainable tourism practices. Include an analysis of the beneficial effects two sustainable tourism projects have had on the local community and environment.		

Sustainable Initiative	Choose a sustainable initiative that could be introduced in the UAE. Discuss ways in which the chosen initiative demonstrates a commitment to sustainable development and its necessity within the UAE and region. Identify any barriers to your chosen initiative. Research, discuss and suggest implementable ideas to enable this initiative to grow. Include an analysis of the beneficial impact of this initiative on the local community and environment.		

It is noteworthy that although the 'Ethics in Social Psychology' assignment guidelines instructed students to research one psychological case, thereby pointing to a Case Study genre, these assignments were also classified in the current study as a Problem Question genre. This is due to Case Studies requiring a large amount of contextual data enabling interpretation of a complex situation, whereas Problem Questions provide a simpler alternative to Case Studies as they enable students to focus on a single issue without the distraction of a large amount of additional authentic data while still applying discipline-specific standard models of analysis to the situation under study (Nesi & Gardner, 2012).

Assignment 3 (Table 4.5) was classified as belonging to the 'Research Report' genre family due to its 'Introduction - Method - Results - Discussion' structure (corresponding to Swalesian moves) and its focus on both empirical research and exploration of relevant literature relating to the issue under investigation (Nesi & Gardner, 2012). The collected Research Reports focused on a wide range of issues selected by individual students, relating to the assignment's topic of barriers to sustainability focusing on various sustainable development initiatives (e.g., thermostats, wind turbines, recycling, pollution reduction, solar energy or LED lights).

Table 4.5: Assignment 3 genre classification

Assignment topic/title	Assignment brief extracts	Writing genre features identified in student texts and brief	Assigned writing genre
Managing Barriers to Sustainability	Include the following: 1. Literature Review: Explore the subject of barriers to sustainable development. Explore how these barriers have affected sustainable development initiatives, people's behaviour and decisions, and how such barriers could be prevented in future. 2. Primary Research: Provide details of how the sample survey was conducted by choosing an appropriate sample size and sampling techniques. Conduct primary research by consulting the relevant respondents. Present the statistics and analyse the data that has been collected. Consider how you will present your data (e.g. tables, graphs, bar charts). 3. Draft Strategy: Based on the literature review and primary research, draw out a conclusion about how various barriers may impact your chosen sustainable development initiative. Outline a strategy for how you would manage the identified barriers.	Genre-based Introduction-Method-Results-Discussion headings Issues are explored with reference to relevant published literature and also through empirical research	Research Report

All titles under Assignment 4 (Table 4.6) displayed the characteristics of 'Discussion Essays' with a 'issue - alternative arguments - final position' structure (Nesi & Gardner, 2012). This is reflected in the collected assignments in the inclusion of alternative positions relating to the issue under discussion before reaching a position based on the evidence discussed.

Table 4.6: Assignment 4 genre classification

Assignment topic/title	Assignment brief extracts	Writing genre features identified in student texts and brief	Assigned writing genre
Social Media and Crime	Discuss the question: *Does social media encourage crime*? Clearly outline both positive and negative influences of social media in relation to crime. Directly answer the question, demonstrating a clear line of argument and drawing a firm conclusion from your findings. Your conclusion should clearly state what your position is on the impact of social media.	Entertainment of alternative positions on an issue before reaching a position or thesis that reconciles them in line with the evidence discussed; these alternative positions might be pros and cons, or alternative claims. Presentation of an issue followed by arguments that can be developed to support different positions on the issue – weighing up or evaluation of evidence to ascertain how strong it is in relation to a position on an issue, leading to: Issue – Alternative arguments – Final position structure.	Discussion Essay
Social Media for Academic Purposes	Discuss the question: *Could students' usage of online social networking for academic purposes have a beneficial or detrimental effect on their experience at university*? Clearly outline both positive and negative influences of social networking when used in a higher education context. Directly answer the question, demonstrating a clear line of argument and drawing a firm conclusion from your findings. Your conclusion should clearly state what your position is on the impact of social media.		
Surveillance Society	Answer one of the following questions: *The benefits of a surveillance society outweigh its detriments; In a surveillance society privacy is an illusion*; *Surveillance does more harm than good; The Surveillance Society is a necessity in the modern world.* Discuss arguments both for and against the statement. Support these arguments with evidence from secondary research. Directly answer the question, drawing a firm conclusion from your findings.		

The above classification of the students' assignments into writing genres forms eleven sub-corpora, grouped topically (Table 4.7).

Table 4.7: Overview of sub-corpora

Assignment & Writing genre	Sub-corpus & Topic	No. of texts (campuses)	Size (tokens)
Assignment 1: Exposition Essays	**E1_LS_42** Learning styles	42 (UK, MRU)	33,789
	E1_MI_38 Multiple intelligences	38 (UK, MRU)	35,326
	E1_FN_113 Fake news	113 (UAE)	77,163
Assignment 2: Problem Questions	**PQ_BOS_48** Business organisational structure	48 (UK, UAE)	80,957
	PQ_ESP_73 Ethics in social psychology	73 (UAE)	118,907
	PQ_ST_59 Sustainable tourism	59 (UK, MRU)	110,660
	PQ_SI_13 Sustainable initiatives	13 (UAE)	21,605
Assignment 3: Research Reports	**RR_193** Managing barriers to sustainability	193 (UK, UAE, MRU)	348,779
Assignment 4: Discussion Essays	**E2_SMC_42** Social media and crime	42 (UK, MRU)	55,815
	E2_SMA_38 Social media for academic purposes	38 (UK, MRU)	51,355
	E2_SS_113 Surveillance society	113 (UAE)	132,991

Table 4.7 shows that the individual sub-corpora were assigned a code. These codes identify the assignment's writing genre, topic and number of texts. For example, E2_SS_113 indicates that it is the second of the two essays (E2) on the topic of *Surveillance Society* (SS) completed by 113 students. The word count of the sub-corpora included in Table 4.7 was generated by AntWordProfiler (Anthony, 2013). This corpus analytical tool excludes punctuation marks from the word count and counts contracted forms (e.g. *isn't*) and possessives (e.g. *Kolb's*) as two words. It is noteworthy that both contracted forms as well as possessives were found to be rare in the students' texts due to the academic nature of their writing; hence, the word count is not expected to be distorted by contracted forms and possessives being counted as two words. This is important to mention as what counts as a token (3.2.1) (i.e. tokenisation) differs from tool to tool; thus, details about the selected tool and how the token count was obtained should be provided (Brezina, 2018).

Since "[t]he exploitation of learner corpus data critically depends on the design of the learner corpus" (Díaz-Negrillo & Thompson, 2013, p. 21), the different sub-corpora in Table 4.7 representing three genre families (Essays, Problem Questions, and Research Reports) and four writing genres (Exposition Essays, Discussion Essays, Problem Questions, Research Reports) will assist in addressing RQ1 (i.e. What are the density

and diversity of academic vocabulary in international foundation students' assessed academic writing across writing genres?). This is enabled owing to the corpus containing data produced in accordance with varied task types (i.e. writing genres), which lends itself to the investigation of the influence of genre in learner language use (Díaz-Negrillo & Thompson, 2013). For the purpose of RQ2 (i.e. To what extent do the density and diversity of academic vocabulary in international foundation students' assessed academic writing develop over an academic year?), only texts belonging to the same genre family (i.e. essays) will be used. RQ3 (i.e. What are the students' perceptions of the contributing factors impacting the acquisition, deployment and development of academic vocabulary?) will be addressed by collecting interview and survey data, discussed next.

4.3.5 Interview and survey data

The interview data were collected in line with the *writing as situated activity* writing paradigm (4.2.2) with the aim of exploring the students' experiences with academic writing during the IFP, particularly their perceptions of the main contributing factors that impacted their acquisition, deployment and development of academic vocabulary (RQ3). Interviews were selected as one of the methods of obtaining data due to their common use in writing research "to learn more about attitudes to writing, about teaching and learning" (Hyland, 2016b, p. 118).

To enable extensive follow-up of the participants' responses, a semi-structured format of interviews was selected as its loose set of guidelines ensures flexibility, which enables the participants to discuss their interpretations and perspectives (Hyland, 2016b). Since academic vocabulary is a key aspect of academic writing style (3.2.3), the interviews were intended to gain insights into the students' experiences with academic writing in general as well as with academic vocabulary in particular. The interviews thus revolved around themes (Appendix 5) exploring various aspects surrounding the students' academic writing on the IFP, such as the students' writing strategies, encountered difficulties, their perceptions of the factors that had contributed to the acquisition and deployment of academic vocabulary in their writing, or perceived changes in their academic writing over the duration of the IFP.

In order to enable the students to reflect on their experience with academic writing during the IFP, the interviews were conducted at the end of the IFP with students

who self-selected to participate in the interview. This resulted in 14 interviews across the three campuses with 4 participants from the UK-based campus, with whom face-to-face interviews were conducted, and Skype interviews with 5 students at each of the two overseas campuses (i.e. 5 with the UAE-based participants and 5 with students in Mauritius).

However, despite the usefulness of interviews in terms of gaining insights into the participants' experiences and perspectives, there are several factors that needed to be acknowledged as they could potentially impact the participants' accounts. One of these relates to the issue of reactivity; that is, the effects of the researcher on the nature of the collected data (Hammersley & Atkinson, 2007), resulting from the fact that I was the UK-participants' tutor on the Researching and Presenting module (4.3.1). The issue of reactivity, however, is likely to have been greatly reduced by the fact that I was not involved in the delivery or assessment of the modules from which assignments were collected for this study. It is, therefore, not believed that there was a major influence on the collected interview data as a result of my positioning as a researcher. In addition, minimising the problem of reactivity is not always a prime consideration provided that the researcher understands how their presence may have shaped the collected interview data, which should be interpreted accordingly (Hammersley & Atkinson, 2007).

Another aspect of interviews that needed to be taken into consideration was the influence that the wording of the interview questions may have on the participants' responses, as emphasised by Kvale (2007, p. 63): "the kind of knowledge produced in the interview depends to a considerable extent on the wording of the questions, which should be in line with the purpose of an interview study." This highlights two points: first, since the aim of the interview was to gain insights into the students' experiences and perceptions, the questions were mainly open-ended; second, the participants' ability to discuss the issue under study was to be achieved by ensuring that the wording of the questions was accessible to the participants by keeping the questions as simple as possible.

A further limitation of interview data, which was difficult to address, relates to the students' lack of awareness and recall of the potential factors that may have played a contributory role in the acquisition, deployment and development of their academic vocabulary on the IFP. Further, due to the timing of the interviews (which

had taken place before the textual analysis was conducted) it was not possible to ask questions relating to specific points of interest arising from the textual analysis. For this reason, although motivated by the academic literacies approach to investigating student writing from the writers' perspectives (2.3), in this study it was not possible to utilise the 'talk around text' technique commonly used in this tradition to gain insights into academic writers' writing practices.

The practical difficulties with conducting Skype interviews with participants based at the overseas campuses also ought to be noted. Due to frequent issues with the internet connection at the Mauritius campus, despite several attempts, Skype calls could often not be established or completed. As for the UAE-based participants, it was only possible to conduct a Skype interview after they had returned to their home country on completion of the IFP due to audio and video applications blocked in the UAE. This meant that very few students were willing to participate in an interview as it coincided with their holidays. Consequently, out of the total of 22 overseas students who agreed to an interview (15 UAE-based and 7 Mauritius-based students), only ten (5 students from each campus) completed an interview.

These practical difficulties concerning the two overseas campuses were recognised prior to conducting interviews and were partially addressed by distributing an online survey (Appendix 6) comprising themes similar to those used in interviews. However, as many questions as possible were designed as multiple-choice questions to increase the completion rate. To achieve responses as accurate as possible, a number of criteria had to be observed when constructing the survey. Among these was the use of clear and unambiguous questions and an inclusion of a range of possible answers which were exhaustive and mutually exclusive in cases of multiple choice questions. The questions were also intended to be precise in asking for the required information and worded in a way that would not suggest what answer was expected (Sapsford, 2007). Nevertheless, one potential issue with this method of data collection relates to not knowing to what extent the participants' responses are genuine. Similar to the interviews, the questionnaires were intended to gain insights not only into academic vocabulary, but also into the students' overall experiences with academic writing on the IFP due to academic vocabulary being an indispensable aspect of academic writing. The survey was completed by 51 students out of the 168 students based overseas (i.e. a 30% response rate).

Despite the above-discussed potential issues, both interview and survey data assisted in addressing RQ3 as they provided useful insights into the students' experiences with academic writing. They also complemented the textual analysis by providing additional understanding of the students' deployment of academic vocabulary.

4.4. Data Analysis

This section provides an account of the processes involved in the collected data analysis. It first focuses on how the textual data were processed to enable a textual analysis (4.4.1), then describes the analytical procedures involved in the textual analysis relating to RQ1 (4.4.2) and RQ2 (4.4.3). This is followed by an overview of the interview and survey data analysis addressing RQ3 (4.4.4).

4.4.1 Textual data processing

Prior to textual analysis of the students' assignments (Table 4.2 in 4.3.4), it was deemed necessary to modify the texts in a number of ways to achieve more accurate results by preserving only the text produced by the students. Therefore, the following amendments were made to all collected assignments: assignments briefs, list of references / bibliography, headers and footers (typically containing the students' name and module number), and end-of-sentence citations (i.e. non-integral) were removed. The rationale behind this was that these are not part of the running text. Sentence-prominent (i.e. integral) citations were preserved in order to avoid distortion of the sentence structure, but additional bibliographic details (e.g. the year of publication and page numbers) were removed. This was intended to achieve a more accurate count of the words representing the continuous prose produced by the students to ensure more reliable measures of academic vocabulary. This was seen as particularly important during the analytical procedures where the density and diversity of academic vocabulary (4.4.2) were calculated in relation to the overall word count.

In addition to the above, appendices, tables and figures including captions were also removed in Problem Questions and Research Reports as most of these were taken from external sources, meaning that the language did not reflect the learners' language production. This amendment is in line with the approach to these textual elements in the BAWE corpus (2.2.4), where tables and figures were not considered part of the running text and were thus excluded. This is considered necessary for

locating phenomena within a text because different parts of texts are of unequal importance and the running texts may be regarded as the text proper (Nesi et al., 2005). In addition, tables and figures are generally not a suitable format to be used with corpus analytical tools. Further, hyphenated words were amended by removing the hyphen (e.g. in-depth -> indepth, high-quality -> highquality, long-term -> longterm). This amendment was necessary in order to ensure that hyphenated words appearing in the AVL (3.2.4) were detected by the selected tool (4.4.2) which does not treat hyphens as a part of a word. All the above-outlined amendments to the collected assignments were intended to ensure accurate measures of academic vocabulary counts.

In all writing genres, direct quotes were preserved as they were found to represent very small proportions of texts. Headings and sub-headings (found primarily in Problem Questions and Research Reports) were also preserved as they are an integral feature of these writing genres. It is also noteworthy that any errors (e.g. misspelt words or incorrect use of word forms) were not corrected in order to preserve the authenticity of the learners' writing. An overview of the above-outlined amendments is provided in Table 4.8.

Table 4.8: Overview of textual amendments

Textual amendments	
Deleted	**Preserved**
• assignment briefs • headers and footers • tables and figures • hyphens in hyphenated words • end-of-sentence citations • list of references / bibliography • appendices	• direct quotes • heading and subheadings • errors (e.g. spelling mistakes, word forms) • sentence-prominent citations (except year and page numbers)

Following these textual amendments, the files were saved (using the same code as assigned to individual students discussed in 4.3.2) in a plain text format using Unicode UTF-8 character encoding to comply with the selected tools' settings (4.4.2). This enabled the textual analysis (RQ1 and RQ2) of the student writing, discussed next.

4.4.2 Data analysis: academic vocabulary across writing genres (RQ1)

In order to address RQ1, the New Academic Vocabulary List (AVL) (Gardner & Davies, 2014) (3.2.4 & 4.2.1) was selected for identification of academic vocabulary items in the students' written assignments. First, however, the AVL was modified in a number

of ways: 1) *disproportionately,* which appears in the AVL twice, was corrected to one entry; 2) UK spellings were added to US variations (e.g. *formalise, organisation, endeavour*); 3) spelling variations were added to words with alternative spelling (e.g. *archaeological* was added to *archeological*, *palaeolithic* was added to *paleolithic, adviser* was added to *advisor*). These amendments resulted in a list containing 3,113 lemmas.

Since the AVL's organising principle is based on lemmas, the collected texts were lemmatised using TagAnt (Anthony, 2015). This tool was selected for its open accessibility, ease of use and ability to lemmatise a large number of files which are subsequently automatically saved in the original folder. Based on these lemmatised texts, the academic vocabulary items were then identified using AntWordProfiler (Anthony, 2013). The suitability of this freeware for this study lies in its ability to analyse the entire corpus at once as well as in batches. Next, although by default this tool analyses texts on the basis of the AWL and GSL (3.2.4), it is possible to upload alternative word lists, such as the AVL selected for this study. A further advantage is in the tool's output, which records the percentage of the academic vocabulary types (unique academic words) and tokens (all academic words) (3.2.1) identified on the basis of the selected academic word list uploaded to the tool. Also, this output can conveniently be saved in an Excel spreadsheet, which enables further customised sorting and other relevant operations.

The identified academic vocabulary items were then checked for their part of speech to ensure that only academic vocabulary lemmas belonging to grammatical classes specified in the AVL were included. This was important due to the differences in meaning that some of these identified academic vocabulary can carry without grammatical identification, for example: *bridge* (noun or verb), *joint* (adjective or noun), *mean* (noun, adjective or verb), *prompt* (verb or adjective), *relative* (noun or adjective), *state* (noun or verb), *stem* (noun or verb), *type* (noun or verb). These academic vocabulary items display one word class only in the AVL (e.g. *bridge* as a verb, *mean* as a noun). This assessment was intended to prevent distortion of frequencies by vocabulary items occurring with high frequencies but not being classified as core academic items in the AVL. For instance, prior to this assessment for grammatical information, *class* (which can be used as a noun and verb but is included as a verb in the AVL) was one of the most frequently deployed academic

vocabulary items in sub-corpora E1_LS_42 and E1_MI_38 as a noun, which would lead to an inaccurate frequency count.

Nation (2001) notes that there are two ways of measuring academic vocabulary in texts: by considering the coverage that academic vocabulary accounts for and by counting the number of academic vocabulary types, lemmas or word families. Accordingly, the extracted academic vocabulary items were measured in terms of: a) coverage referring to the proportion of academic vocabulary in relation to all words (henceforth 'density'), which was calculated as a percentage of academic tokens per all tokens; b) the number of academic types in relation to all types referring to the range or variety of academic vocabulary (henceforth 'diversity'), calculated as a percentage of academic types per all types. It should also be noted that in the context of this study, *academic types* refer to the different academic lemmas as opposed to "a unique word form" (Brezina, 2018, p. 39) belonging to the same stem. This is because the academic vocabulary items were identified using lemmatised texts, which did not contain inflected word forms. Both density and diversity of academic vocabulary were based on the output of AntWordProfiler (Anthony, 2013).

The analysis of academic vocabulary from the perspective of the density and diversity of academic vocabulary was intended to provide insights into the composition of the students' texts in terms of the proportion of academic vocabulary items in relation to all words (i.e. the density of academic vocabulary) and the repertoire of academic vocabulary items (i.e. the diversity of academic vocabulary) in the different writing genres under study (RQ1) as well as how the density and diversity of academic vocabulary change over the duration of the foundation programme (RQ2).

To explore the density and diversity of academic vocabulary, the Shapiro-Wilk test of normality (Shapiro & Wilk, 1965) was conducted first to establish whether the data were normally distributed. This test was selected over the alternative Kolmogorov-Smirnov test as it is more appropriate for sample sizes up to 50 and is also reported to be reliable with samples as large as 2,000. Thus, the Shapiro-Wilk test was used to test for normality in all sub-corpora and showed that all data were normally distributed, including sub-corpora containing outliers, which were investigated and found to be genuine values.

Next, the following measures were obtained: mean (M) to measure central tendency of the data; standard deviation (SD), coefficient of variation (CV) and range to

measure the overall dispersion of the values; and 95% confidence interval (CI) for the mean to obtain "an estimation of the true value of a statistical measure" (Brezina, 2018, p. 13) indicating that we can be 95% confident that the mean falls between the identified values. In this study, the t-statistics was used to calculate the CI rather than the z-statistics since the sample's (not the population's) SD was used to calculate the CI. The mean and the range are displayed in a boxplot, which was selected for visualisation purposes over other visual data representations as it "provides much more information about the data than a simple barchart" (Brezina, 2018, p. 23), such as the position of the median in relation to the mean as well as the interquartile range showing where 50% of the values lie. For the purpose of addressing RQ1, only the values of the mean, range and outliers are displayed in the boxplots as these are sufficient for the analysis and subsequent interpretation and discussion of the findings.

The rationale behind employing two different dispersion measures (i.e. SD and CV) is to assess the variation within a sub-corpus (using SD) as well as across sub-corpora, whereby the CV was used for comparison of variation between sub-corpora. This is because the CV is a standardised measure and can thus be used to compare variation across sub-corpora with different means (Brezina, 2018). In addition, a one-way ANOVA was used to determine whether there were statistically significant differences in the density and diversity of academic vocabulary between the different writing genres as well as among the sub-corpora of the same writing genre. This was reported using the p value complemented by Cohen's *d* effect size (Cohen, 1988). In this study, the conventional cut-off point of 0.05 ($p<.05$) is adopted for the p-value and 0.3 for a small effect size, 0.5 for a medium effect size and 0.8 for a large effect size (Brezina, 2018). The one-way ANOVA was followed by a t-test with Bonferroni correction to establish between which sub-corpora the differences lie, also reported using the p-value and Cohen's *d* effect size.

It is noteworthy that although most sub-corpora comprised texts produced by students based at different campuses, these were not analysed separately by location as the IFP followed the same curriculum and assessment procedures across the three campuses. In addition, in the vast majority of cases the participants were not permanent residents or citizens in their place of study (4.3.3).

Following the above-described procedures, both density and diversity of academic vocabulary were further explored in individual sub-corpora, described next.

Density of academic vocabulary

The density of academic vocabulary was further explored by investigating how it occurred; that is, which academic vocabulary items formed the biggest proportion of academic vocabulary in the sub-corpora. To gain insights into the proportion of academic vocabulary items in relation to all academic tokens, percentages were used and academic vocabulary items forming at least 1% of all academic tokens in each sub-corpus were extracted. Some of these academic vocabulary items were found to appear in the assignment brief and were marked with an asterisk (*).

In addition to providing information about the proportion of the extracted academic vocabulary items in relation to all academic vocabulary tokens, the following information is provided: the range (i.e. the number of texts in which the identified academic vocabulary items appeared), raw or absolute frequency (i.e. the actual count of occurrences of the academic vocabulary items in the sub-corpus), and normalised or relative frequency (i.e. the absolute frequency of the item normalised to a common base). The normalisation of frequencies to a rational common base is seen as important where corpora of various sizes are being compared (Brezina, 2018; McEnery et al., 2006). It is suggested that the relative frequency is normalised to a basis similar in size to the corpus or its sub-corpora. Therefore, the identified academic vocabulary items were normalised to a common base of 10,000 words, widely regarded as an appropriate rate for smaller corpora (Brezina, 2018). Both raw and normalised frequencies are included since “relative frequencies should never be used to hide the absolute frequency but should be reported together with absolute frequencies” (Brezina, 2018, p. 44).

The most frequently deployed academic vocabulary items forming at least 1% of all academic tokens were further examined with the assistance of AntConc (Anthony, 2018) using its concordance plot function, which provided information about the location of the academic vocabulary items under investigation. This was followed by a close reading of selected texts in order to obtain a more comprehensive insight into the usage of the academic vocabulary items. Where available and relevant, these insights from the textual analysis were complemented by interview and / or survey data and extracts from the assignment briefs to obtain further understanding of the

deployment of the most frequently used academic vocabulary items in the various writing genres under study.

Diversity of academic vocabulary

The diversity of academic vocabulary was approached from the perspective of the identified academic vocabulary types distribution among students to explore how the diversity of academic vocabulary occurred; that is, what proportion of students shared the same academic vocabulary items. To consider the different sizes of the sub-corpora in terms of the number of texts they comprised (with each text representing one student), this was calculated using the percentage of students (i.e. texts) in each sub-corpus (i.e. ≥75%, ≥50%, ≥25%, <25% of students, and individual students) in relation to the percentage of academic vocabulary types (i.e. what percentage of academic vocabulary types was shared by what percentage of students). This was intended to provide insights into the extent of core academic vocabulary shared by the majority (i.e. at least 50%) of students. The rationale behind selecting academic vocabulary items used by the majority of students (as opposed to selecting a frequency-based approach with a frequency cut-off point) was the aim to arrive at typicality of usage among the students, avoiding identification of academic vocabulary items overused by a small number of students.

Since "frequency statistics alone do not tell us everything" (McCarthy & Carter, 1997, p. 25), the academic vocabulary items deployed by at least 50% of students were further explored for their functions in their co-textual environment following Hyland's (2008) functional categories of academic clusters (loosely based on Halliday's linguistic metafunctions of ideational, interpersonal and textual, discussed in 2.2.1), which were also found suitable for classification of individual academic vocabulary items in the present study. These categories are: *research-oriented* (referring to academic vocabulary describing the writers' activities and experiences of the real world) further sub-divided into location, procedure, quantification, description, and topic; *participant-oriented* (focusing on the reader or writer of the text) including stance and engagement features; and *text-oriented* (referring to discourse organisers) containing transition, resultative, structuring and framing signals.

These categories were adopted in this study as follows: research-oriented location items referred to places (e.g. *university*), procedure items were primarily action

words (e.g. *apply*, *improve*, *activity*), description items were mainly adjectives or nouns providing more specific information about the item under discussion (e.g. *sustainable*, *beneficial*, *society*), and topic vocabulary items were closely related to the assignment topic and often contained in the assignment title (e.g. *theory*, *harmful*, *structure*); text-oriented transition signals referred to vocabulary items expressing additive or contrastive links (e.g. *furthermore*, *however*), resultative items expressed causative relationships (e.g. *therefore*, *consequently*), framing items specified limiting conditions (e.g. *within*, *both*), and structuring items were those organising stretches of discourse or had the function of directing the reader elsewhere in the text (e.g. *firstly*, *finally*, *above*) and were also used to classify headings / sub-headings (e.g. *introduction*, *literature review*) as well as vocabulary used as signposts (e.g. *state*, *argue*, *aim*, *purpose*) and referring to Tables and Figures. It is noteworthy that no research-oriented quantification or participant-oriented vocabulary items were identified among those academic vocabulary items deployed by at least 50% of the students.

The functional categories of the academic vocabulary used by the majority of the students were examined with the assistance of a concordancing programme AntConc (Anthony, 2018) in order to achieve an accurate assessment of the items' functions and meanings in their co-textual environment. This is important as some identified academic vocabulary carry different meanings in different contexts and thereby display a different function dependent on their context. A case in point is *general* used as an adjective with a research-oriented descriptive function, and as an adverb (i.e. in *general*) displaying a text-oriented framing function, identified in several sub-corpora.

However, it has to be pointed out that "[w]hile these classifications are sufficiently broad to minimize the possibility of overlaps between categories, no system is entirely watertight" (Hyland, 2008, p. 49). Accordingly, it was also found that some identified academic vocabulary displayed more than one function in the same sub-corpus. In such cases, an examination of the concordance lines was conducted so as to "verify the most appropriate category based on the typical use" of the item (Gray, 2016, p. 37) to avoid assigning more than one category to each lexical item. In addition, some of the identified academic vocabulary items were found to be prevalent in the assignment briefs. However, academic vocabulary items drawn from

the assignment briefs were not assigned an individual functional category as they often belonged to one of the above categories. Instead, they were marked with an asterisk (*). For a clearer presentation of the identified functional categories, the codes in Table 4.9 were used.

Table 4.9: Academic vocabulary functional categories codes

Function	Code	Function	Code
Research-oriented:	RO	Text-oriented:	TO
location	RO-loc	transition	TO-trans
procedure	RO-proc	resultative	TO-res
description	RO-desc	structuring	TO-str
topic	RO-top	framing	TO-fram

(adapted from Hyland, 2008)

In order to ensure an accurate classification of the functional categories, the functions of the academic vocabulary were assessed twice at a two-month interval. The intra-rater reliability is reported using raw agreement, referring to "a metric, often expressed as a percentage, which provides the proportion of agreement cases in all cases" (Brezina, 2018, p. 89). This metric is considered suitable for the purpose of intra-rater reliability in this context, despite the fact that it is proposed to rate inter-rater reliability (Brezina, 2018). In this study, the raw agreement was as follows: 81.08%, 93.75% and 95% for the three Exposition Essays; 91.78%, 98.18%, 94.74% and 97.53% for the four Problem Question assignments; 100% for the Research Reports; 100%, 97.14% and 93.55% for the Discussion Essays. This can be considered acceptable since an agreement of 80% and above is desirable (Brezina, 2018).

In addition to the functional categories, similar to the information provided for the density of academic vocabulary, the following information is provided about the extracted vocabulary items: the range, raw and normalise frequency.

4.4.3 Data analysis: academic vocabulary development (RQ2)

RQ2 explores the development of academic vocabulary in the students' assessed academic writing over the duration of the IFP delivered over a period of six months (4.3.1). It is important to note, however, that contemporary writing development research represents a complex scene as definitions of what constitutes 'development' in writing vary and "have been in a state of flux over the past fifty years" (Camp, 2012, p. 93).

Due to the complexities in defining this concept, writing development can thus refer to various phenomena. "[C]ontemporary applications of developmental theory nonetheless lead us to some shared assumptions about growth that can guide our efforts to foreground development in the assessment of writing" (Camp, 2012, pp. 93–94), which is often "inferred from the observation of changes in concrete samples of L2 production collected at different times, such as essays or other writing samples in the case of written production" (Bulté & Housen, 2014, p. 46).

In this study, therefore, the notion of development refers to the growth in density (i.e. the proportion of academic vocabulary in relation to all words) and diversity (i.e. the range or variety) of academic vocabulary over the duration of the IFP, and was investigated by making comparisons between the Exposition Essays (E1) and Discussion Essays (E2) in terms of increasing or decreasing density and diversity of academic vocabulary.

The rationale behind selecting these two writing genres for the purpose of addressing RQ2 was two-fold: a) they represent the first and final written assignment on the IFP (Table 4.7 in 4.3.4) completed by individual students (Table 4.1 in 4.3.4) with the E1 thus serving as a baseline; b) although representing two different writing genres, the Exposition and Discussion Essays belong to the same Essay genre family (4.3.4) (Nesi & Gardner, 2012), meaning that they share the same social purpose and can thus be regarded as suitable for comparative purposes for examining academic writing development in this context.

To visualise the differences between E1 and E2, findings are displayed in boxplots showing the average (i.e. mean) density and diversity of academic vocabulary and the range. For the purpose of addressing RQ2, similar to RQ1 (4.4.2), only the mean, range and outliers are displayed in the boxplots as these values are considered sufficient for the analysis, interpretation and subsequent discussion of the findings.

Further, to establish whether the differences between E1 and E2 were statistically significant, a paired sample t-test was conducted. This test is commonly used to compare the mean values of a selected linguistic variable between two groups (Brezina, 2018). Due to its most important assumptions (i.e. independence of observation and normal distribution) being met (Brezina, 2018), this test was considered appropriate in this context. The result of the t-test is reported using the t-value together with a p-value (using a cut-off point of $p<.05$) and Cohen's *d* effect

size (0.3 for a small effect size, 0.5 for a medium effect size and 0.8 for a large effect size). The effect size measure (i.e. the size of the difference between the two groups) is important to include so as to evaluate the t-test in standardised terms (Brezina, 2018).

Next, the number of students who showed an increase as well as decrease in both density and diversity of academic vocabulary was calculated together with the percentage points of the increase and / or decrease. The students showing the highest increase and decrease in the density and diversity of academic vocabulary and students with an increase and decrease closest to the average (i.e. mean) were further investigated with reference to the students' texts complemented by interview and survey data where available. The focus on the students showing the highest, lowest and average increase and decrease in both density and diversity of academic vocabulary was intended to provide initial insights into the phenomenon under investigation, and prompted further exploration of the potential factors impacting the changes in the density and diversity of academic vocabulary.

This led to an investigation into the effect of the academic vocabulary contained in the assignment brief by measuring the changes in these academic vocabulary items between E1 and E2 in relation to all academic tokens (density) and types (diversity). This was calculated as a percentage (rounded to whole numbers) of academic vocabulary items contained in the assignment brief per all academic tokens (density) and types (diversity) in both E1 and E2. Next, the potential impact of the assignment topic on the changes in the density and diversity of academic vocabulary was explored by analysing paired sub-corpora separately with the aim of establishing whether certain topics prompted a denser and more diverse usage of academic vocabulary than others. Specifically, the following sub-corpora were formed by a paired sample of students: E1_LS_42 and E2_SMC_42; E1_MI_38 and E2_SMA_38; E1_FN_113 and E2_SS_113 (Table 4.7 in 4.3.4). As above, a paired sample t-test was used to establish whether the differences in the density and diversity of academic vocabulary between the paired sub-corpora were statistically significant.

In addition, the diversity of academic vocabulary was further explored by investigating the changes in the size of the core academic vocabulary by comparing the distribution of the academic vocabulary types shared by ≥75%, ≥50%, ≥25%, <25% of students, and individual students. This was followed by comparing the changes in

the function of the academic vocabulary types deployed by the majority (i.e. at least 50%) of students between E1 and E2, based on Hyland's (2008) categorisation (Table 4.9 in 4.4.2).

When comparing the functional categories of the identified academic vocabulary in E1 and E2, some of the identified academic vocabulary items were also found to display the function of metadiscourse, referring to an umbrella term encompassing a range of cohesive and interpersonal features which assist in relating text to its context by explicitly organising the writer's stance towards its readers or the text's content. Metadiscourse thus serves as a link between a text and its context by addressing the expectations that readers have of the texts with regard to certain forms of interaction and engagement (Hyland, 2005). Metadiscourse can therefore be seen as a defining feature of successful writing, meaning that an awareness of metadiscourse features and the ability to use these features effectively are important in the development of student writing. Based on its function in a text, Hyland (2005) proposes an interpersonal model of metadiscourse, which recognises two aspects of interaction: *the interactive dimension* concerned with helping to guide the reader through the text and *the interactional dimension* involving the reader in the text, both comprising further sub-categories. The interactive dimension includes 'transition markers' (e.g. *thus*), 'frame markers' (e.g. *finally*), 'endophoric markers' (e.g. *noted above*), 'evidentials' (e.g. *according to*), and 'code glosses' (e.g. *namely*). The interactional dimension comprises 'hedges' (e.g. *perhaps*), 'boosters' (e.g. *definitely*), 'attitude markers' (e.g. *surprisingly*), 'engagement markers' (e.g. *note that*), and 'self mentions' (e.g. *we*). The identified academic vocabulary items serving as metadiscourse were related to this categorisation.

4.4.4 Data analysis: Students' perceptions (RQ3)

RQ3 investigates the IFP students' perceptions of the main contributing factors impacting their academic vocabulary. Specifically, the focus is on productive academic vocabulary (3.2.2) deployed in the students' written assignments. For the purpose of exploring the students' perspectives, questionnaire and interview data were utilised, both collected at the end of the IFP so as to enable the students to reflect on their experiences throughout the IFP (4.3.5). Since academic vocabulary is a key aspect of academic writing style (3.2.3), the questionnaire as well as interviews were intended to provide insights not only into the students' experiences with

academic vocabulary, but also into academic writing on the IFP in general. Therefore, not all questions contained in the survey and interview schedule led to informing RQ3. From the survey (Appendix 6), question 14 (below) was the main source of information in addressing RQ3.

14. Which of the following writing strategies did you use when completing your essays and reports on the IFP? (tick as many as you like)

- paraphrasing (i.e. using own words)
- summarising (i.e. extracting main points)
- translating
- using direct quotes (i.e. using the wording of the source text)
- using vocabulary from texts
- using vocabulary from assignments briefs/instructions
- using vocabulary from teaching materials
- other…

Regarding the interview data (Appendix 5), question 8 (below) served as the main source of insights into the students' perceptions regarding their academic vocabulary.

8) Vocabulary development

Q: (How) do you think your academic vocabulary used in written assignments has developed over the course of the academic year?

Q: What do you think has contributed to this?

(prompts, if needed: taught sessions, independent learning, reading)

The survey data were analysed and reported quantitatively using the proportion and percentage (rounded to whole numbers) of responses relating to the relevant aspect of the question (e.g., 22/51, 43% of students). For the analysis of the interview data, thematic analysis was adopted as "it offers an accessible and theoretically flexible approach to analysing qualitative data" (Braun & Clarke, 2006, p. 77), which can potentially result in a rich, detailed and complex account of data. In line with thematic analysis, the interview data were examined so as to identify various themes, referring to something that the researcher considers important in relation to the research question (Braun & Clarke, 2006). Depending on the aims of the qualitative analysis, there are two approaches to a qualitative data analysis utilising thematic analysis: a rich description of the entire data set, or a detailed account of one particular aspect

of the data (Braun & Clarke, 2006). Considering the specificity of RQ3 focusing on the students' perspectives of and experiences with academic vocabulary, the latter approach to qualitative analysis was adopted with the aim of providing a detailed account of the students' experiences with academic vocabulary on the IFP.

A further decision in conducting a thematic analysis concerns the inductive (i.e. 'bottom-up') versus deductive (i.e. 'top-down') approach to themes identification. While the inductive approach strongly links the themes to the data themselves and the identified themes may have little relation to the questions asked, the deductive approach is driven by the researcher's interest in a particular aspect. Consequently, the inductive thematic analysis is likely to result in a rich description of the data set overall, while the deductive thematic analysis often leads to a more detailed description of some aspect of the data (Braun & Clarke, 2006). Accordingly, this study followed the deductive approach to thematic analysis.

Next, the level at which themes are identified needs to be considered. This can be done either at a semantic (i.e. explicit) level or at a latent (i.e. interpretative) level. The semantic approach to thematic analysis identifies the themes within the explicit meanings of the data and does not focus on anything beyond the participants' response. In contrast, the latent approach explores the underlying ideas, assumptions and conceptualisation by going beyond the semantic content of the data (Braun & Clarke, 2006). For the purpose of addressing RQ3, the semantic level was selected as a suitable approach to identifying themes as the analysis at this level typically progresses from description of the data (which have been organised according to semantic content) to summary and interpretation of the significance of the patterns.

In sum, as regards the approaches to thematic analysis for the purpose of addressing RQ3, a deductive thematic analysis describing a particular aspect of the data set was employed focusing on a semantic level of analysis. These decisions were followed by six phases of thematic analysis, as proposed by Braun and Clarke (2006). First, the collected interview data were transcribed and subsequently checked against the audio recordings to ensure that the transcripts were as true to the participants' verbal accounts as possible (e.g. making sure that punctuation added did not alter the meaning). Second, having read and become familiar with the interview data, an overview of ideas was generated reflecting the various relevant aspects of interest.

This led to the production of initial codes relating to the features at the semantic level of the data that were found to be of relevance to RQ3. Considering the relatively small number of participants (i.e. 14 students) and the duration of the interviews (i.e. approximately half an hour), the coding was done manually. Examples of the initial codes include 'teaching', 'readings', 'journal articles', 'books', 'feedback', or 'group work', for instance. After all data had been coded, the next step involved identifying themes, whereby the various initial codes were grouped into potential overarching themes. During this phase, the codes 'journal articles', 'books' and 'readings' were grouped together to form the theme of 'reading sources', for example.

The next phase involved a refinement of the themes identified in the previous phase, resulting in some themes being collapsed together. For instance, the theme of 'reading sources' was merged with the theme of 'reference sources' to form the theme of 'appropriate sources'. This phase was followed by further defining and refining the themes, referring to identification of the essence of each theme as well as the themes overall, together with determining what aspect of the interview data each of the themes captures. This resulted in the identified themes forming two overarching themes relating to the students' perceptions of the main contributing factors impacting their productive academic vocabulary. One of these overarching themes was interpreted as relating to the sources of new academic vocabulary forming the overarching theme of 'acquisition of academic vocabulary'. Another overarching theme related to the perceived factors that impacted the improvement of the deployment of academic vocabulary in the students' written production and was labelled 'development of academic vocabulary'. It has to be pointed out, however, that in the context of RQ3, the meaning of 'development' differs from that in RQ2 where 'development' refers to changes (i.e. increase in the density and diversity of academic vocabulary) based on a textual analysis, whereas in RQ3 'development' refers to the improvement in the usage of productive academic vocabulary as perceived by the students.

The final phase of thematic analysis is producing a write-up of the thematic analysis, which should include sufficient evidence of the identified themes. Hence, the identified themes and sub-themes were supported with data extracts in the form of participants' quotations obtained from the interview data demonstrating the relevant aspects of the themes.

These two overarching themes identified on the basis of the interview data were complemented by the questionnaire data (question 14), which related to the students' writing strategies and provided insights into the students usage of academic vocabulary. This theme was thus labelled 'deployment of academic vocabulary'.

In sum, the survey and interview data led to three overarching themes relating to the students perceptions of the contributing factors that impacted their academic vocabulary on the IFP: 'acquisition of academic vocabulary', 'deployment of academic vocabulary', and 'development of academic vocabulary'.

4.5 Methodology: summary

This chapter has discussed the theoretical frameworks and approaches relating to the investigation of writing (4.2) which are particularly relevant to the phenomenon under investigation in the present study and how these informed the data collection procedures (4.3) and subsequent analysis (4.4).

The above-discussed writing paradigms and their underpinning ontological and epistemological assumptions (4.2) are reflected in this study by conducting a corpus-based analysis of the students' assessed academic writing (i.e. their summative written assignments) representing the writing product relating to the *writing as completed activity* writing paradigm (4.2.1) to address RQ1 and RQ2. As one of the criticisms of learner corpus data is that the producers of the data are no longer available to the researcher, meaning that it is not possible to further examine language production, this limitation can be overcome by drawing on various other data sources (referred to as triangulation), which provide further information about the phenomenon under study (Díaz-Negrillo & Thompson, 2013). Hence, the corpus-based textual analysis was complemented by participant interviews and surveys with the aim of gaining insights into how different individuals perceived the contexts surrounding their assessed academic writing relating to the *writing as situated activity* writing paradigm (4.2.2). This was intended to address RQ3 exploring the students' experiences with academic writing on the IFP and their perceptions of the main contributing factors that impacted the deployment and development of academic vocabulary items in their assessed academic writing.

It is important to note that although the deployment of academic vocabulary is often linked to academic achievement (1.2), this study did not investigate whether there was a link between the students' awarded grades and the deployment of academic vocabulary in their written production. The reason for this is two-fold: firstly, the assessment criteria with which the students had been provided prior to completion of the written assignments referred to the use of a suitable writing style in general without explicitly emphasising the usage of academic vocabulary in the students' written production; secondly, this is an exploratory study focusing on the learners' production and perspectives rather than approaching this from the perspective of a marker.

The collected data were intended to address this study's research questions, as follows:

- RQ1 (What are the density and diversity of academic vocabulary in international foundation students' assessed academic writing across writing genres?) was addressed by exploring academic vocabulary items in the four identified writing genres forming the basis for this study by conducting a corpus-based analysis of the students' assessed academic writing
- RQ2 (To what extent do the density and diversity of academic vocabulary in international foundation students' assessed academic writing develop over an academic year?) explored the changes in the deployment of academic vocabulary items over the period of an academic year by making comparisons between the Exposition and Discussion Essays submitted at the beginning and end of the IFP, also utilising corpus-based methodology
- RQ3 (What are the students' perceptions of the contributing factors impacting the acquisition, deployment and development of their academic vocabulary?) was addressed by individual semi-structured interviews and online surveys focusing on the students' experiences with academic writing and the main contributing factors impacting the acquisition, deployment and development of academic vocabulary.

The theoretical frameworks and approaches as well as the data collection and analytical procedures described in this chapter are believed to be suitable in generating insights into the phenomenon under investigation in the present study as they were intended to generate results in response to all three RQs.

The next three chapters present the results relating to the three RQs accompanied by discussion.

Chapter 5: Academic vocabulary across writing genres

5.1 Introduction

This chapter presents the results and discussion relating to the deployment of academic vocabulary in the four writing genres of assessed academic writing (i.e. Exposition Essays, Problem Questions, Research Reports, and Discussion Essays) which formed the corpus of the student assessed academic writing investigated in this study (4.3.4). Table 5.1 provides a summary of the various sub-corpora representing the individual genres together with the assignment topics, number of texts forming each sub-corpus as well as the students' location and information relating to the size of the sub-corpora in terms of running words (i.e. tokens), average number of tokens per text, the required word length as stated in the assignment brief, the raw number of academic tokens and the percentage of academic tokens per all tokens, the raw number of types as well as the raw number of academic types and percentage of academic types per all types.

In this chapter, the two ways of measuring academic vocabulary (density and diversity) are focused on separately (4.4.2). First, the results and discussion of the density of academic vocabulary in the four writing genres under investigation are focussed on, followed by the results and discussion of the diversity. Specifically, an overview of the density of academic vocabulary across all four writing genres is provided (5.2), followed by a detailed analysis and presentation of results relating to individual writing genres and their sub-corpora (5.2.1 – 5.2.4). Next, the results generated by the analysis of the density of academic vocabulary in all four writing genres are discussed (5.3). The following section focuses on the diversity of academic vocabulary across the four writing genres (5.4) together with a detailed analysis and presentation of results relating to individual writing genres and the sub-corpora they comprise (5.4.1 – 5.4.4). These results are discussed in the next section (5.5). Throughout this chapter, visualisations are used to aid the presentation of results and their subsequent discussion. These are primarily in the form of boxplots, tables and figures containing extracts from the students' texts. This chapter concludes with a summary of the main findings (5.6).

Table 5.1: Sub-corpora overview

Sub-corpus (Assignment topic)	No. of texts (campus)	Size (tokens)	Average tokens per text	Required word length[10]	No. of ac. tokens (% per tokens)	No. of types	No. of ac. types (% per types)
Exposition Essays							
E1_LS_42 (Learning styles)	42 (UK, MRU)	33,789	800	720 (+/- 10%)	5,649 (16.72%)	2,569	777 (30.25%)
E1_MI_38 (Multiple intelligences)	38 (UK, MRU)	35,326	930	720 (+/- 10%)	4,962 (14.05%)	2,656	701 (26.39%)
E1_FN_113 (Fake news)	113 (UAE)	77,163	680	650 – 750	8,972 (11.63%)	4,437	874 (19.7%)
Problem Questions							
PQ_BOS_48 (Business organisational structure)	48 (UK, UAE)	80,957	1,700	1,700 (+/- 10%)	15,102 (18.65%)	4,522	1,028 (22.73%)
PQ_ESP_73 (Ethics in social psychology)	73 (UAE)	118,907	1,630	1,700 (+/- 10%)	21,078 (17.73%)	4,796	1,043 (21.75%)
PQ_ST_59 (Sustainable tourism)	59 (UK, MRU)	110,660	1,900	1,700 (+/- 10%)	21,428 (19.36%)	5,323	1,142 (21.45%)
PQ_SI_13 (Sustainable initiatives)	13 (UAE)	21,605	1,660	1,700 (+/- 10%)	3,502 (16.21%)	2,110	529 (25.07%)
Research Report							
RR_193 (Managing barriers to sustainability)	193 (UK, MRU, UAE)	348,779	1,800	1,700 (+/- 10%)	66,453 (19.05%)	8,597	1,579 (18.37%)
Discussion Essays							
E2_SMC_42 (Social media and crime)	42 (UK, MRU)	55,815	1,330	1,350 (+/- 10%)	9,054 (16.22%)	4,269	879 (20.59%)
E2_SMA_38 (Social media for academic purposes)	38 (UK, MRU)	51,355	1,350	1,350 (+/- 10%)	10,896 (21.22%)	3,338	851 (25.49%)
E2_SS_113 (Surveillance society)	113 (UAE)	132,991	1,180	1,200 - 1,500	19,252 (14.48%)	5,746	1,183 (20.59%)

[10] as stated in the assignment brief in the Assessment Handbook

5.2 Density of academic vocabulary: results

This section provides the results relating to the density of academic vocabulary in the four writing genres under investigation. An overview of the density of academic vocabulary in the four writing genres is given in Figure 5.1, followed by an analysis of individual sub-corpora forming these different writing genres (Table 5.1).

Figure 5.1: Density of academic vocabulary across writing genres

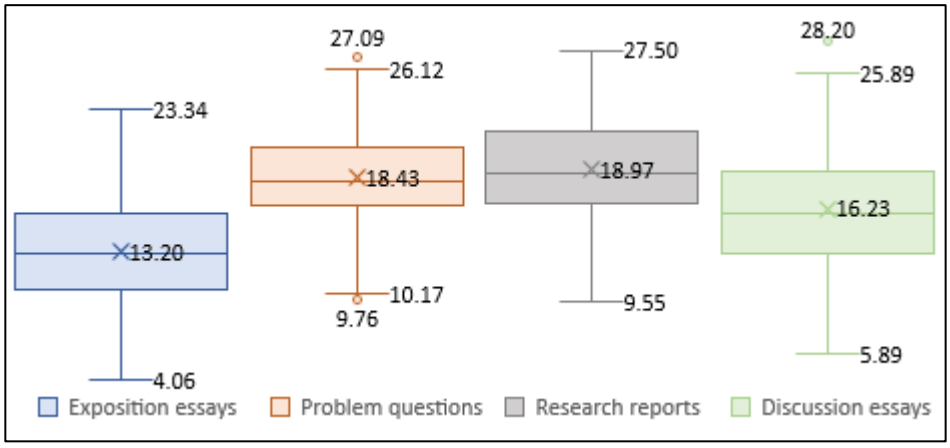


Figure 5.1 displays the mean average of the density of academic vocabulary across the four writing genres (i.e. 13.2% in Exposition Essays, 18.43% in Problem Questions, 18.97% in Research Reports, 16.23% in Discussion Essays) together with the range showing the overall dispersion of the values (i.e. the distance between the smallest and the largest value). In addition, outliers (i.e. extreme values that are far from the other values) can be seen in the Problem Questions sub-corpus (i.e. 9.76% and 27.09%) and in the Discussion Essays sub-corpus (i.e. 28.2%).

As Figure 5.1 shows, the density of academic vocabulary varied across the four writing genres with the Exposition Essays showing the lowest and the Research Reports the highest density of academic vocabulary. The differences among the four writing genres were statistically significant (as confirmed by a one-way ANOVA: f=98.97, p<.05). Statistically significant differences with large and medium effect sizes were also found between all pairs of the four sub-corpora except between Problem Questions and Research Reports (as confirmed by a post-hoc t-test with Bonferroni correction)[11]. Since the majority of the writing genres were represented by several

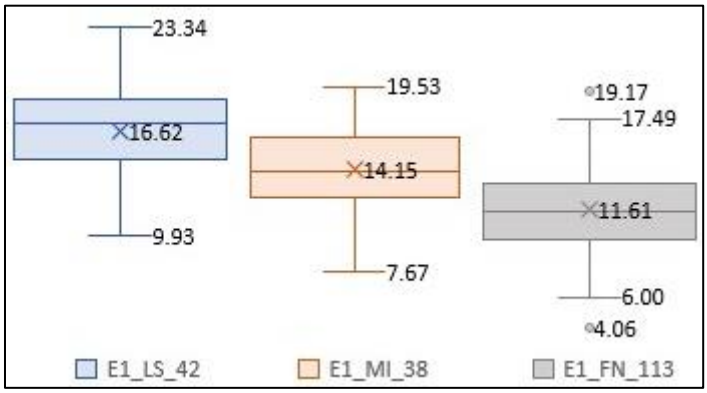
[11] A post-hoc t-test with Bonferroni correction results: Exposition Essays and Problem Questions (t=14.03, p<0.01, d=1.57), Exposition Essays and Research Reports (t=15.49, p<0.01, d=1.62), Exposition Essays and Discussion Essays (t=8.14, p<0.01, d=0.77), Problem Questions and Research Reports (t=1.46, p=0.87, d=0.16), Problem Questions and Discussion Essays (t=5.89, p<0.01 d=0.59), Research Reports and Discussion Essays (t=7.35, p<0.01, d=0.69).

sub-corpora (Table 5.1 above), the density of academic vocabulary was investigated in more detail with the aim of exploring the similarities and differences in the density of academic vocabulary not only across but also within the writing genres.

5.2.1 Assignment 1 Exposition Essays

This section presents results relating to the writing genre of Exposition Essays comprising the following three sub-corpora: E1_LS_42 (Learning Styles); E1_MI_38 (Multiple Intelligences); and E1_FN_113 (Fake News) (Figure 5.2).

Figure 5.2: Density of academic vocabulary (Exposition Essays)



E1_LS_42: SD=2.89, CV=17.4%, 95% CI=[15.72, 17.52]
E1_MI_38: SD=3.02, CV=21.32%, 95% CI=[13.16, 15.14]
E1_FN_113: SD=2.83, CV=24.39%, 95% CI=[11.08, 12.14]

The density of academic vocabulary in the three sub-corpora of Exposition Essays varied considerably (Figure 5.2) and ranged from 4.06% (E1_FN_113) to 23.34% (E1_LS_42), with the sub-corpus E1_FN_113 showing the lowest (M=11.61%) and E1_LS_42 showing the highest (M=16.62%) density of academic vocabulary. The differences between the three sub-corpora were statistically significant (as confirmed by a one-way ANOVA: f=48.76, p<.00001). Statistically significant differences were also found between all pairs of sub-corpora of Exposition Essays[12], indicating that factors other than the writing genre impacted the density of academic vocabulary in Exposition Essays, further explored next.

Sub-corpus E1_FN_113 showed the highest internal variation (CV=24.39%) (also evident from two outliers) (Figure 5.2). This is interesting considering that sub-corpus E1_FN_113 was the only sub-corpus of Exposition Essays containing essays produced

[12] A post-hoc t-test with Bonferroni correction results: E1_LS_42 and E1_MI_38 (t=3.83, p<.01, d=0.84); E1_LS_42 and E1_FN_113 (t=9.61, p<.01, d=1.75); and E1_MI_38 and E1_FN_113 (t=4.7, p<.01, d=0.87).

by students in one location (i.e. the UAE), while the other two sub-corpora comprised two groups of students (based in the UK and Mauritius) (Table 5.1). Hence, this suggests that factors other than the place of study impacted the density of academic vocabulary in the Exposition Essays. Also, sub-corpus E1_FN_113 contained significantly more texts (i.e. 113) than the other two sub-corpora (i.e. 42 and 38), which could have been a factor impacting internal variation.

The above hence points to a relationship between a higher internal variation and a lower density of academic vocabulary (i.e. the density of academic vocabulary decreased with an increasing internal variation), and between the number of texts and internal variation (i.e. internal variation increased with the increasing number of texts). Thus, the different number of texts in each sub-corpus of Exposition Essays seemed to correlate with the internal variation, which may in turn have impacted the density of academic vocabulary. The relationship between the internal variation, number of texts and density of academic vocabulary was hence investigated in the other writing genres (5.2.2 - 5.2.4) to establish whether it is a recurring finding across the four writing genres under study.

The next section explores the density of academic vocabulary in Exposition Essays in more depth by investigating how it was achieved. Tables 5.2 - 5.4 show academic vocabulary items forming at least 1% of all academic vocabulary tokens in Exposition Essays.

Table 5.2: Academic vocabulary forming ≥ 1% of all academic tokens (E1_LS_42)

Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens	Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens
*experience**	42	259	76.65	4.58	*process*	33	69	20.42	1.22
*theory**	36	164	48.54	2.9	*strategy*	23	67	19.83	1.19
study	35	117	34.63	2.07	*knowledge*	33	60	17.76	1.06
group	32	114	33.74	2.02	*category*	19	59	17.46	1.04
information	33	89	26.34	1.58	*identify**	26	58	17.17	1.03
*type**	29	72	21.31	1.27	*subject*	25	58	17.17	1.03

*academic vocabulary item in the assignment brief

Table 5.3: Academic vocabulary forming ≥ 1% of all academic tokens (E1_MI_38)

Academic lemma	Range	Raw freq.	Norm freq.	% of ac tokens	Academic lemma	Range	Raw freq.	Norm freq.	% of ac tokens
*theory**	36	147	41.61	2.96	*individual*	24	68	19.25	1.37
*multiple**	34	113	31.99	2.28	*subject*	26	66	18.68	1.33
*experience**	35	91	25.76	1.83	*group*	24	60	16.98	1.21
develop	25	75	21.23	1.51	*possess*	23	59	16.7	1.19
improve	23	73	20.66	1.47	*apply*	24	53	15	1.07
type	22	72	20.38	1.45	*knowledge*	25	51	14.44	1.03
logical	28	71	20.1	1.43	*information*	27	50	14.15	1.01
*example**	28	68	19.25	1.37	*university**	27	50	14.15	1.01

*academic vocabulary item in the assignment brief

Table 5.4: Academic vocabulary forming ≥ 1% of all academic tokens (E1_FN_113)

Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens	Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens
social	101	414	53.65	4.61	*result*	56	128	16.59	1.43
*harmful**	109	410	53.13	4.57	*impact*	58	122	15.81	1.36
information	98	360	46.65	4.01	*individual*	48	118	15.29	1.32
affect	78	211	27.34	2.35	*example*	68	113	14.64	1.26
article	60	194	25.14	2.16	*therefore*	60	109	14.13	1.21
society	63	187	24.23	2.08	*influence*	59	107	13.87	1.19
state	70	165	21.38	1.84	*discuss**	76	105	13.61	1.17
effect	71	133	17.24	1.48	*report*	59	98	12.7	1.09
source	64	129	16.72	1.44	*publish*	49	97	12.57	1.08

*academic vocabulary item in the assignment brief

Tables 5.2 - 5.4 contain 12, 16 and 18 academic vocabulary items (respectively) and show that academic vocabulary items forming the highest proportion of academic tokens appeared in the assignment brief (marked *), including those in the assignment question (e.g. 'Is Fake News *Harmful*' in E1_FN_113) (Table 5.5).

Table 5.5: Exposition Essays assignment briefs extracts

Assignment topic & sub-corpus	Assignment brief extracts
Learning Styles E1_LS_42	Draw upon your own past experiences of learning. Analyse your own learning style and how it relates to the characteristics of the learner types identified in Kolb's theory. Show how the learning theory relates to you as a learner. Apply Kolb's theory to yourself (it is not sufficient to repeat the exact wording of Kolb's theory). Your points should be illustrated using examples from your past learning experiences. Illustrate points with experiences from past learning.
Multiple Intelligences E1_MI_38	Explain the theory of multiple intelligences and how this relates to learning. Include examples of your learning practice illustrating how your approach to learning fits with Gardner's theory. Illustrate points with experiences from past learning.
Fake News E1_FN_113	Is fake news harmful? Discuss arguments either for or against the motion.

The impact of the assignment brief on the density of academic vocabulary was explored further by considering all academic tokens in relation to the assignment brief (Table 5.6).

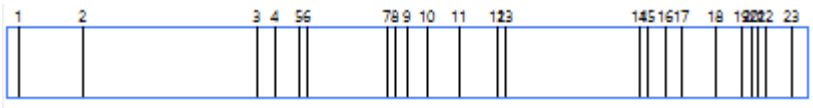
Table 5.6: Impact of assignment brief on the density of academic vocabulary (Exposition Essays)

Sub-corpus	No. of tokens	No. of ac. lemmas in brief	% of ac. lemmas from brief per ac. tokens	% of ac. lemmas from brief per all tokens
E1_LS_42	165	26	17.63%	2.95%
E1_MI_38	156	22	16.61%	2.33%
E1_FN_113	110	18	8.83%	1.03%

Table 5.6 shows that a higher number of academic vocabulary items contained in the brief was reflected in a higher proportion of these expressions forming academic tokens as well as all tokens in the students' texts. This points to the effect that the assignment brief may have had on the deployment of academic vocabulary items in the students' Exposition Essays, meaning that the assignment brief may have been an important source of academic vocabulary for students. This is supported by other researchers (e.g., Flowerdew, 2003; Milton, 2001) who also found that some of the lexical items investigated in their studies of student writing could be traced back to the assignment guidelines (3.2.2), further discussed in 5.3.1. Therefore, "[t]he influence of related texts … should not be underestimated in the analysis of learner corpora" (Flowerdew, 2003, p. 497). The most frequently deployed academic vocabulary items drawn from the assignment briefs of Exposition Essays presented in Tables 5.2 - 5.4 above were thus investigated in more detail.

In Table 5.2 (above), *experience* was the only academic lemma used by all students. The lemma *experience* appeared three times in the E1_LS_42 brief: "…you should draw upon your own past *experiences*"; "Your points should be illustrated using examples from your past learning *experiences*"; "Illustrate points with *experiences* from past learning" (Table 5.5 above). The 42 students in sub-corpus E1_LS_42 used *experience* with varying frequencies (M=6) up to 23 occurrences per student (Student UK16/17/D; Figure 5.3).

Figure 5.3: Usage of *experience* Student UK16/17/D (E1_LS_42)



The concordance plot in Figure 5.3 shows the location of *experience.* It can be seen that Student UK16/17/D used *experience* throughout the essay. This is further shown in an extract from this Student's essay (Figure 5.4).

Figure 5.4: Student UK16/17/D essay extract (E1_LS_42)

In addition, an accommodator learns from hands-on experiences could help me have more experiences. In particular, when I was in high school, students have to attend all the classes, sit there and listen to what the teachers said about all the subjects. It might be difficult for an accommodator person like me, because I was just sitting there and listening to teachers without practising. Instead of listening to the teachers, an accommodator who learns from hands-on experiences prefer to get involved in experiences, in order to develop their knowledge as well as experiences. Furthermore, I have accommodator tendencies, I prefer to solve problems and have my own experiences, instead of listening to other people. Also, nobody can do everything correctly for the first time, so a learner who learns from hands-on experiences will keep attempting the best that they can to solve the problems. …

Accordingly, as an accommodator having new experiences I could achieve more goals in my future. To put it more simply, new experiences help me more successful in future and set up my schedule more clearly and succeed beyond my expectations.

To sum up, this essay described my past learning experiences and compared with Kolb's four learning styles as I am an accommodator likes doing things, learning from hands-on experiences and getting involved in new experiences I prefer having new experiences by myself. Besides, new experiences help me set up a better schedules and achieve more goals in my life even hurt myself or failed. As a result, having new experiences could help me reach a good result and succeed in my future.

Figure 5.4 shows a rather frequent usage of *experience* in the Student's essay. However, it should be expected that lexical items from the brief were incorporated in students' writing as this is a clear way for the student to show that they were addressing the assignment instructions (Milton, 2001). Nevertheless, the dense and repeated usage of *experience* in short sections of the text in Figure 5.4 could be regarded as inappropriate lexical overuse. A possible explanation for this might be under-developed academic writing skills or a lack of lexical repertoire of some L2 writers. The latter was also found by Staples and Reppen (2016) whose study showed that L2 writers relied more on repetition of vocabulary than L1 writers (2.2.3). A further possible reason for this lexical overuse might be the Student's deliberate repeated use of this vocabulary item as a strategy to address the assignment instructions due to the emphasis on the discussion of the learning experiences in the brief.

Further, it also ought to be acknowledged that repetition is common in some languages, meaning that lexical overuse in some students' assignments may have

been impacted by their L1 conventions. Liu and Zhang (2012), for instance, observed ineffective use of simple repetition (i.e. “an occasion when a lexical item that has already occurred in a text is repeated with no big alternation”, p. 112) in Chinese EFL writing with the majority of Chinese L2 learners of English having a tendency to use words to repeat the meaning and form at sentence or textual level. Repeating the same word can thus be seen as a characteristic feature of Chinese EFL writing practices, where repetition is not only tolerated by Chinese speakers, but it is also encouraged in educational settings: “Chinese EFL learners are even encouraged to repeat the same thing for sentence coherence, textual coherence and for paragraph development, since they are seldom taught about the English writing convention” (Liu & Zhang, 2012, p. 125). This is in stark contrast to English writing conventions, where repetition is sparingly used with synonyms and ways of restating or paraphrasing ideas being frequently-used writing strategies (Liu & Zhang, 2012). Therefore, the dense and repeated usage of certain academic vocabulary items in the IFP students’ writing could also stem from the students’ L1 writing conventions, which they have transferred to their L2 academic writing.

Further insights were gained from interview data available for Student UK16/17/D. In the interview, the Student stated that one of the difficulties encountered with academic writing was the use of own words and paraphrasing, which may have been a key factor in the frequent repetition of *experience* drawn from the brief. This was investigated further in the Student’s essay by considering all academic tokens in relation to those drawn from the assignment brief, and it was found that out of the 75 academic tokens used by the Student, 32 were drawn from the assignment brief (i.e. 43% of academic tokens). It may thus be hypothesised that students who are less proficient or confident in using their own words and / or paraphrasing techniques (including the use of synonyms) tended to rely and draw on the assignment brief as a source of vocabulary to incorporate in their writing. This finding is consistent with that of Milton (2001) who noted greater reliance of less proficient students on the assignment prompt (further discussed in 5.3.1).

In sub-corpus E1_MI_38 (Table 5.3 above), the three most frequently used academic vocabulary items (i.e. *theory*, *multiple*, *experience*) were contained in the assignment brief with *theory* and *multiple* not only confirming the impact of the assignment brief (noted above), but also pointing to the effect of the assignment topic (*Theory* of

multiple intelligences). The important role that the topic played in student text compositions was confirmed by some students in the surveys. Student MRU17/18/J, for instance, considered "having in depth research on the topic at hand" an important aspect of good academic writing at university, similar to Student MRU18/M who also thought that being "well organized on a topic" is important in academic written production.

The interview data provided further support to the students' perceptions of the importance of the topic in their written assignments:

> "I tend to go off topic at some point…so staying more relevant to the topic is something I look forward to achieving" (Student UAE18/K).
>
> "You need to focus on the topic…if you don't impress the lecturer by sticking to the topic you might fail" (Student MRU18/F).

The effect of the topic on the deployment of vocabulary may be expected, however, as "[e]ach text has its own topic vocabulary which occurs because of the message the text is trying to convey" (Nation, 2001, p. 208), and has also been noted by others (e.g. Olinghouse & Wilson, 2013) (3.2.5) (further discussed in 5.3.2).

In terms of the link between less proficient students and the reliance on academic vocabulary items contained in the assignment brief (as hypothesised above), survey data were available for eight students from this sub-corpus (out of the total of 38 students forming this sub-corpus). Six of them stated that they found the use of appropriate vocabulary or own words difficult. In one of these students' essays (Student MRU18/R), the academic vocabulary items drawn from the assignment brief formed nearly one quarter of all academic tokens (23%) and around 3% of all tokens in their text, illustrated in Figure 5.5.

Figure 5.5: Student MRU18/R essay extract (E1_MI_38)

This essay will examine Gardner's abstraction of multiple intelligences and how they relate to me. In addition, I will discuss what type of learner I am. My past experiences will then be scrutinised and how I will develop my intelligences for university use.

Musical intelligence is the ability to replicate, identify and reflect on music which is generally illustrated by Musicians, Vocalists and Sensitive listeners. According to Gardner Musical intelligence includes the appreciation of music, not only that but the skill of composing and performing music. I define to musical intelligence because I am very talented in Music, I also appreciate music because it helps me relate to certain things like the past or the present. Music helps me reflect as a person and also helps me stay focused. In high school I excelled in the subject music, not only in composing and performing but also in analysing, researching and reflecting other musicians works of all genres.

In conclusion, mastering Gardner's theory of multiple intelligences can enlighten ways of learning. By identifying my ways of learning, it can help me build up my learning practices. Moreover, Identifying and developing my intelligences is very significant in university because it will help me be outstanding in my studies. Assimilating Garner's theory of multiple intelligences will assist me to develop my three dominating intelligences in order to be excellent in my studies at University.

Hence, the analysis of sub-corpus E1_MI_38 for the usage of academic vocabulary items contained in the assignment brief generated findings consistent with those for sub-corpus E1_LS_42, supporting the hypothesis that the assignment brief was an important source of vocabulary for some students.

As for sub-corpus E1_FN_113, out of the 18 academic vocabulary items forming more than 1% of all academic tokens (Table 5.4 above), *harmful* was contained in the assignment brief as well as in the assignment question / title (i.e. Is fake news *harmful*?). Similar to sub-corpus E1_MI_38, this points to the impact of both the assignment brief as well as topic on the density of academic vocabulary (further discussed in 5.3.1 and 5.3.2). Since *harmful* was contained in the assignment question, it is interesting that not all students seem to have answered the essay question directly by using *harmful*; 109 out of 113 used *harmful* in their essay with frequencies up to 13 instances (M=4) (Student UAE17/18/H, Figure 5.6).

Figure 5.6: Usage of *harmful* Student UAE17/18/H (E1_FN_113)

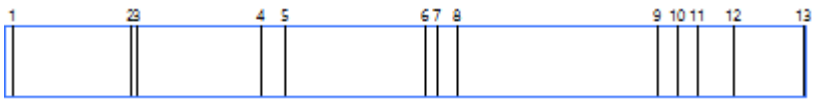


Figure 5.6 shows that Student UAE17/18/H used *harmful* throughout the essay. A further examination of the Student's essays showed that s/he answered the essay question by making numerous references to fake news being *harmful* (Figure 5.7).

Figure 5.7: Student UAE17/18/H essay extract (E1_FN_113)

This essay will discuss whether fake news is harmful. … This essay will conclude that fake news is harmful.

Firstly, fake news is harmful as it leads to the misinforming of the general population … Therefore, fake news is harmful as it misinforms the people of false things …

Furthermore, fake news is harmful as it also affects people mentally … The mental effects of fake news on people all seem to be negative and harmful to their psychological health. Hence, fake news is harmful as it causes emotional distress to people involved.

Another reason why fake news may be harmful is that it leads to the consequence of the mistrust in the media. … Therefore, fake news is harmful as it causes mistrust in the media, social media and newspapers alike.

In conclusion, fake news is harmful. It keeps circulating until legal action is taken and has a lot of negative effects that are harmful to people. … Because of the many negative effects it has, it can be considered to be harmful. … This essay comes to conclude that fake news is harmful.

Similar to the results obtained from sub-corpora E1_LS_42 and E1_MI_38 (above), Figure 5.7 illustrates a repeated use of an academic vocabulary item drawn from the assignment brief and title. Survey data were available for 35 students forming this sub-corpus (out of the total of 113 students forming this sub-corpus), out of whom 13 reported encountering difficulties relating to appropriate vocabulary usage with seven of them stating that they drew on the assignment brief as a writing strategy when completing their written assignments. However, in total, 18 students reported having drawn on the assignment brief during the process of completing their written assignments. This, hence, suggests that drawing on the assignment brief may also have been a strategy employed by students in general so as to address the assignments requirements, not just by students experiencing difficulties with the usage of appropriate vocabulary.

To explore what strategies were employed by the students who did not draw on the academic vocabulary item contained in the assignment brief and title (i.e. *harmful*), the four students' essays not containing *harmful* were investigated. A close reading of these essays showed that the students used paraphrasing techniques instead (Figures 5.8 and 5.9).

Figure 5.8: Student UAE17/18/QQ essay extract (E1_FN_113)

Throughout this paper, the harm of fake news is carefully examined and discussed. … Fake news has affected people politically, socially and economically creating negative impressions worldwide and causing harm.

Figure 5.9: Student UAE18/M essay extract (E1_FN_113)

> The purpose behind this article is to feature the various perils of fake news and its negative impacts on the general population. …
>
> Fake news is physically, financially, socially and politically harming to customers and the world at large.

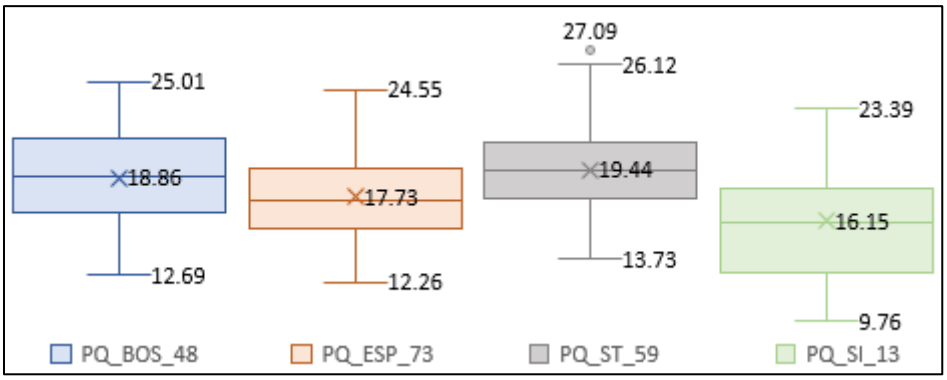
The use of paraphrasing techniques in Student UAE17/18/QQ's essay (Figure 5.8) was supported by questionnaire data from this Student, where s/he selected paraphrasing as one of the strategies used during the process of completing written assignments. This student did not state that s/he drew on the assignment brief or found the use of appropriate vocabulary or own words difficult.

The above analysis of the academic vocabulary items appearing in the assignment briefs of Exposition Essays suggests that the brief was an important source of vocabulary, particularly for students less confident in using their own words and paraphrasing and less familiar with academic vocabulary. These students may have incorporated the vocabulary items from the brief in their writing as a strategy to compensate for the lack of appropriate language required in assessed academic writing at a foundation-level of study, whereas students who did not encounter difficulties relating to academic vocabulary usage tended to use alternative ways of expressing ideas, such as paraphrasing. It is important to acknowledge, however, that these alternatives may not always have been available to students. For instance, the two most frequently used academic vocabulary items in sub-corpus E1_MI_38 (*theory* and *multiple*) related directly to the assignment topic (i.e. Theory of multiple intelligences); hence, it is reasonable to assume that the students would not be expected to use their own words in these instances.

5.2.2 Assignment 2 Problem Questions

This section presents results relating to the writing genre of Problem Questions comprising four sub-corpora: PQ_BOS_48 (Business Organisational Structure); PQ_ESP_73 (Ethics in Social Psychology); PQ_ST_59 (Sustainable Tourism); and PQ_SI_13 (Sustainable Initiatives) (Figure 5.10).

Figure 5.10: Density of academic vocabulary (Problem Questions)



PQ_BOS_48: SD=3.04, CV=16.1%, 95% CI=[17.98, 19.74]
PQ_ESP_73: SD=2.75, CV=15.54%, 95% CI=[17.09, 18.37]
PQ_ST_59: SD=3, CV=15.42%, 95 CI=[18.66, 20.22]
PQ_SI_13: SD=4.23, CV=26.22%, 95% CI=[13.5938, 18.71]

Figure 5.10 shows that the density of academic vocabulary in the four sub-corpora of Problem Questions ranged from 9.76% (PQ_SI_13) to 27.09% (PQ_ST_59), with the sub-corpus PQ_SI_13 showing the lowest (M=16.15%) and PQ_ST_59 showing the highest (M=19.44%) density of academic vocabulary. The differences between the four sub-corpora were statistically significant (as confirmed by a one-way ANOVA: f=6.33, p=.0004), with statistically significant differences also found between several pairs of the sub-corpora[13]. As with Exposition Essays (5.2.1), this suggests that factors other than the writing genre impacted the density of academic vocabulary, further investigated next.

Sub-corpus PQ_SI_13 showed the highest internal variation (CV=26.22%) with the other three sub-corpora showing a similar level of internal variation (CV=15.42% - 16.1%). It is noteworthy that sub-corpus PQ_SI_13 contained significantly fewer texts (i.e. 13) compared to the other sub-corpora (i.e. 48, 59 and 73) and displayed the lowest density of academic vocabulary. Hence, similar to the Exposition Essays, the density of academic vocabulary decreased with an increasing internal variation. Unlike the Exposition Essays, however, internal variation decreased with a larger number of texts.

[13] A post-hoc t-test with Bonferroni correction confirmed statistically significant differences between the following sub-corpora: PQ_BOS_48 and PQ_SI_13 (t=2.87, p<.05, d=0.82), PQ_ESP_73 and PQ_ST_59 (t=3.24, p<.01, d=0.6), and PQ_ST_59 and PQ_SI_13 (t=3.56, p<.01, d=1.01), but not between PQ_BOS_48 and PQ_ESP_73 (t=2.02, p=0.27, d=0.39), PQ_BOS_48 and PQ_ST_59 (t=0.99, p=1.95, d=0.19), and between PQ_ESP_73 and PQ_SI_13 (t=1.74, p=0.5, d=0.44).

The next section explores how the density of academic vocabulary was achieved. Tables 5.7 - 5.10 show academic vocabulary items which formed at least 1% of all academic tokens.

Table 5.7: Academic vocabulary forming ≥ 1% of all academic tokens (PQ_BOS_48)

Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens	Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens
*structure**	47	864	106.72	5.72	*sustainable*	40	179	22.11	1.19
*project**	46	498	61.51	3.3	*product*	32	173	21.37	1.15
organization	40	354	43.73	2.34	*hierarchy*	42	171	21.12	1.13
control	45	283	34.96	1.87	*level*	40	171	21.12	1.13
*environment**	42	250	30.88	1.66	*both**	40	170	21	1.13
environmental	42	198	24.46	1.31	*social*	36	161	19.89	1.07
management	39	198	24.46	1.31	*need*	41	160	19.76	1.06
*beneficial**	31	182	22.48	1.21					

*academic vocabulary item in the assignment brief

Table 5.8: Academic vocabulary forming ≥ 1% of all academic tokens (PQ_ESP_73)

Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens	Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens
experiment	71	2172	182.66	10.3	*researcher*	64	351	29.52	1.67
study	71	1001	84.18	4.75	*psychological*	66	321	27	1.52
*ethical**	73	934	78.55	4.43	*ethic**	69	314	26.41	1.49
*research**	72	688	57.86	3.26	*result*	65	313	26.32	1.48
*guideline**	67	465	39.11	2.21	*subject*	48	273	22.96	1.3
consent	70	385	32.38	1.83	*state*	68	236	19.85	1.12
*social**	68	368	30.95	1.75	*group*	43	229	19.26	1.09
conduct	69	357	30.02	1.69					

*academic vocabulary item in the assignment brief

Table 5.9: Academic vocabulary forming ≥ 1% of all academic tokens (PQ_ST_59)

Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens	Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens
*sustainable**	59	1489	134.56	6.95	*practice**	52	274	24.76	1.28
*environment**	59	427	38.59	1.99	*increase*	54	266	24.04	1.24
impact	57	420	37.95	1.96	*information*	55	237	21.42	1.11
organisation	42	375	33.89	1.75	*sector*	51	224	20.24	1.05
sustainability	50	309	27.92	1.44	*research**	54	221	19.97	1.03
promote	53	297	26.84	1.39	*project**	48	219	19.79	1.02
*development**	53	294	26.57	1.37					

*academic vocabulary item in the assignment brief

Table 5.10: Academic vocabulary forming ≥ 1% of all academic tokens (PQ_SI_13)

Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens	Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens
*initiative**	13	204	94.42	5.83	*barrier**	13	47	21.75	1.34
*sustainable**	12	157	72.67	4.48	*resource*	11	47	21.75	1.34
*benefit**	12	82	37.95	2.34	*farming*	2	43	19.9	1.23
*development**	12	79	36.57	2.26	*increase*	11	42	19.44	1.2
waste	8	79	36.57	2.26	*future*	13	39	18.05	1.11
reduce	13	73	33.79	2.08	*environ-mental*	11	37	17.13	1.06
project	10	63	29.16	1.8	*implemen-tation*	13	37	17.13	1.06
*environment**	11	62	28.7	1.77	*vertical*	1	37	17.13	1.06
*sustainability**	11	48	22.22	1.37	*economic*	11	35	16.2	1

*academic vocabulary item in the assignment brief

Tables 5.7 - 5.10 show that with the exception of the PQ_ESP_73 sub-corpus, the academic vocabulary items forming the biggest proportion of all academic tokens were contained in the assignment brief including those in the assignment title. This supports the effect of the assignment brief and topic on the deployment of academic vocabulary also found in the Exposition Essays (5.2.1), suggesting that this may be a recurring finding across writing genres. The impact of the assignment brief was explored further by considering all academic tokens in relation to the brief (Table 5.11).

Table 5.11: Impact of assignment brief on the density of academic vocabulary (Problem Questions)

Sub-corpus	No. of tokens	No. of ac. lemmas in brief	% of ac. lemmas from brief per ac. tokens	% of ac. lemmas from brief per all tokens
PQ_BOS_48	129	15	15.7%	2.86%
PQ_ESP_73	112	13	15.67%	2.78%
PQ_ST_59	117	20	18.82%	3.64%
PQ_SI_13	159	25	23.81%	3.86%

Table 5.11 shows that in sub-corpus PQ_SI_13, academic vocabulary items drawn from the assignment brief formed the highest proportion of academic tokens (where these academic vocabulary items formed nearly a quarter of all academic vocabulary tokens) as well as all tokens (where they formed nearly 4% of all tokens). It can also be seen that the proportion of academic vocabulary items drawn from the assignment brief per all academic tokens and per all tokens increased with the increasing number of these items contained in the brief. This points to the effect that the brief had on the deployment of academic vocabulary items by students in their

Problem Question assignments; that is, the higher number of academic vocabulary items contained in the brief, the higher proportion of students' texts these vocabulary items formed. Similar to the writing genre of Exposition Essays (5.2.1), this suggests that the assignment brief may have been an important source of academic vocabulary for students. Hence, the academic vocabulary items drawn from the assignment briefs (Tables 5.7 - 5.10) were investigated further to gain insights into the deployment of these vocabulary items in the Problem Question writing genre.

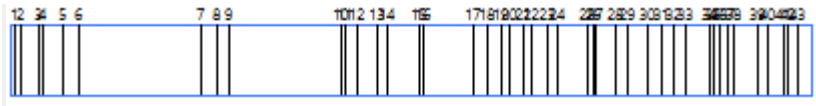
In sub-corpus PQ_BOS_48 (Table 5.7), *structure* was the most frequently used academic vocabulary item. It appeared in the assignment title (Business Organisational *Structure*) and was also the most frequent academic vocabulary item in the assignment brief, used five times (Table 5.12).

Table 5.12: Problem Question assignment brief extract (PQ_BOS_48)

Assignment topic & sub-corpus	**Assignment brief extract**
Business Organisational Structure PQ_BOS_48	Create an organisational structure for a company that will start and deliver a project that will benefit the environment in your country. Research two companies … and compare the organisational structures of both companies. You will use this research … to inform and justify the organisational structure of the company that you are planning. You must justify your structure design … explain why this organisational structure has been chosen …

Structure was used by 47 out of 48 students with frequencies up to 43 uses per student (Student UAE17/18/K) (M=18). An examination of Student UAE17/18/K's text showed that *structure* was used throughout the assignment (Figure 5.11), similar to the results from Exposition Essays (Figure 5.7 in 5.2.1).

Figure 5.11: Usage of *structure* Student UAE17/18/K (PQ_BOS_48)



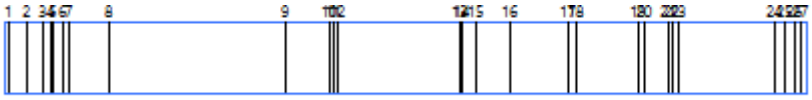
No interview or survey data were available for this Student to gain further insights into the writing strategies employed by this Student when completing written assignments.

It is also noteworthy that the majority of instances of *structure* were in combination with *organisational*, pointing to the effect of the assignment topic (i.e. Business Organisational Structure) on the deployment of academic vocabulary. In addition to the topic effect, *organisational structure* is a common collocation (Cambridge

Dictionary, 2020), which may have contributed to its frequent usage. The topic effect corroborates the findings generated by the Exposition Essays analysed above and further discussed in 5.3.2.

In sub-corpus PQ_ESP_73 (Table 5.8), *ethical* was the only academic vocabulary item contained in the assignment brief deployed by all students and was used up to 27 times per student (Student UAE18/O, Figure 5.12) (M=13).

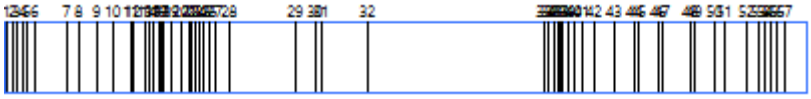
Figure 5.12: Usage of *ethical* Student UAE18/O (PQ_ESP_73)



As Figure 5.12 shows, *ethical* was used throughout the assignment, similar to Figures 5.6 (E1_FN_113) and 5.11 (PQ_BOS_48). The usage of *ethical* also points to the effect of the assignment topic (i.e. Ethics) on the deployment of academic vocabulary, in line with findings from sub-corpus PQ_BOS_48 above.

In sub-corpus PQ_ST_59 (Table 5.9), the most frequently used academic vocabulary item (*sustainable*) was used by all students with frequencies up to 57 (M=25) occurrences per text. An investigation of the Student's text with the highest usage of *sustainable* (i.e. 57 occurrences Student MRU17/18/G) showed that the Student deployed this academic vocabulary item throughout the assignment (Figure 5.13).

Figure 5.13: Usage of *sustainable* Student MRU17/18/G (PQ_ST_59)



Similar to the other two sub-corpora of Problem Questions above, the usage of *sustainable* points to the effect of the assignment topic (i.e. *Sustainable* tourism) on the deployment of academic vocabulary.

As for sub-corpus PQ_SI_13 (Table 5.10), the most frequently deployed academic vocabulary item (*initiative*) was used by all students with frequencies up to 27 occurrences per text (M=16). An investigation of the Student's text with the highest usage of sustainable (i.e. 27 occurrences Student UAE17/18/GG) showed that the Student used *initiative* throughout the assignment (Figure 5.14).

Figure 5.14: Usage of *initiative* Student UAE17/18/GG (PQ_SI_13)

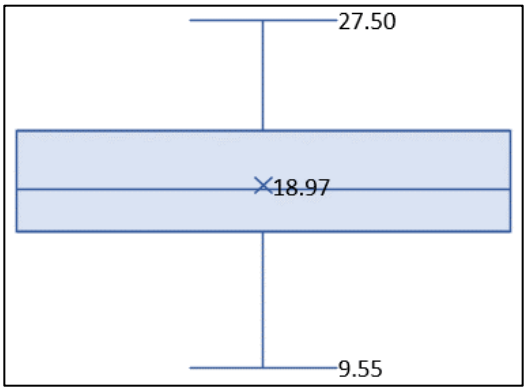
Initiative, together with the second most frequently deployed academic vocabulary item in this sub-corpus (i.e. *sustainable* – also contained in the assignment brief), point to the effect of both the assignment brief and topic (i.e. *Sustainable initiatives*) on the deployment of academic vocabulary, consistent with the other sub-corpora of Problem Questions, further discussed in 5.3.1 and 5.3.2.

The above analysis of the most frequently deployed academic vocabulary items in the Problem Question assignments points to the effect that both the assignment brief and the topic had on the deployment of these vocabulary items in the students' assignments. Similar to the Exposition Essays, the academic vocabulary drawn from the assignment brief offered few alternatives to students. Therefore, it is assumed that students were expected to incorporate these academic vocabulary items in their assignments to some extent to address the assignment instructions. However, the frequencies with which these items were used varied among students, with a considerable number of students over-using these vocabulary items in small sections of their texts.

5.2.3 Assignment 3 Research Report

This sections presents results relating to the writing genre of Research Reports focusing on the topic of 'Managing barriers to sustainability' (Figure 5.15).

Figure 5.15: Density of academic vocabulary (Research Reports)



SD=3.61, CV=19.04%, 95% CI=[18.46, 19.48]

Figure 5.15 shows that the average density of academic vocabulary in Research Reports was 18.97% with an overall range between 9.55% - 27.50% and some internal

variation (CV=19.04%). Table 5.13 below shows how the density of academic vocabulary was achieved in this writing genre.

Table 5.13: Academic vocabulary forming ≥ 1% of all academic tokens (RR_193)

Ac. lemma	Range	Raw freq.	Norm freq.	% of ac. tokens	Ac. lemma	Range	Raw freq.	Norm freq.	% of ac. tokens
*barrier**	191	3242	92.95	4.88	*study**	173	942	27.01	1.42
*sustainable**	183	2384	68.35	3.59	*sustaina-bility**	153	939	26.92	1.41
*initiative**	159	1459	41.83	2.2	*result*	173	784	22.48	1.18
*development**	164	1354	38.82	2.04	*social*	149	766	21.96	1.15
*research**	189	1337	38.33	2.01	*lack*	155	694	19.9	1.04
*survey**	149	1337	38.33	2.01	*method*	135	688	19.73	1.04
figure	125	1125	32.26	1.69					

*academic vocabulary item in the assignment brief

Table 5.13 contains 13 academic vocabulary types forming at least 1% of all academic tokens with 8 (62%) of these items drawn from the assignment brief including those contained in the assignment title (*Managing barriers to sustainability*). As with Exposition Essays (5.2.1) and Problem Questions (5.2.2), this thus points to the effect of the assignment brief and title on the deployment of academic vocabulary in this writing genre.

The impact of the assignment brief was explored further. The assignment brief (which was 455 words in length) contained 39 academic vocabulary lemmas, which were used 19,852 times in total, forming approximately 30% of academic tokens and nearly 6% of all tokens. With regard to the students' reliance on the vocabulary items contained in assignment brief, this points to similarities of Research Reports with the writing genres of Exposition Essays (Table 5.6 in 5.2.1) and Problem Questions (Table 5.11 in 5.2.2); however, the academic vocabulary contained in the assignment brief formed a bigger proportion of the Research Reports than in the two other writing genres analysed thus far.

Table 5.13 above shows that *barrier* formed the biggest proportion of academic vocabulary tokens in the Research Report; it was used 3,242 times by 191 out of 193 students with frequencies up to 59 occurrences per student (Student UK15/16/G, Figure 5.16) (M=17).

Figure 5.16: Usage of *barrier* Student UK15/16/G (RR_193)

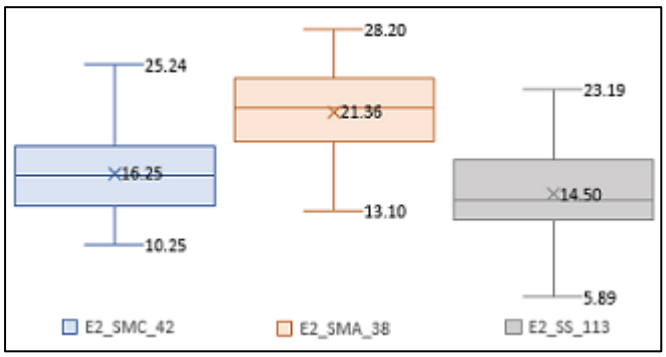
Figure 5.16 shows that the Student used *barrier* throughout the text, in line with the results generated by the Exposition Essays (Figures 5.3 and 5.6) and Problem Questions (Figures 5.11 – 5.14).

An examination of the two students who did not use *barrier* (Students UAE17/18/SS and UAE18/F) showed that Student UAE17/18/SS used alternative expressions such as *challenges of sustainability* and Student UAE18/F failed to address the assignment instructions. Survey data were available for Student UAE17/18/SS, in which the Student stated that s/he used paraphrasing techniques when completing writing assignments and that s/he did not encounter difficulties relating to using appropriate vocabulary or own words. Hence, this result is comparable to that generated by the Exposition Essays and Problem Questions as it also shows that students used two strategies when completing the Research Reports (i.e. drawing on the assignment brief and using paraphrasing techniques), confirming that the assignment brief may have been an important source for many students when completing Research Reports.

5.2.4 Assignment 4 Discussion Essays

This section presents the results relating to the writing genre of Discussion Essays comprising the following three sub-corpora: E2_SMC_42 (Social Media and Crime); E2_SMA_38 (Social Media for Academic Purposes); and E2_SS_113 (Surveillance Society) (Figure 5.17).

Figure 5.17: Density of academic vocabulary (Discussion Essays)



E2_SMC_42: SD=3.21, CV=19.77%, 95% CI=[15.25, 17.25]
E2_SMA_38: SD=3.73, CV=15.79%, 95% CI=[20.25, 22.47]
E2_SS_113: SD=3.43, CV=23.64%, 95% CI=[13.86, 15.14]

Figure 5.17 shows that the density of academic vocabulary in the three sub-corpora of Discussion Essays varied considerably and ranged from 5.89% (E2_SS_113) to 28.20% (E2_SMA_38) with the E2_SMA_38 sub-corpus showing the highest (M=21.36%) and E2_SS_113 showing the lowest (M=14.50%) density of academic vocabulary. A one-way ANOVA confirmed that the differences between the sub-corpora were statistically significant (f=58.83, p<.00001). A post-hoc t-test with Bonferroni correction showed statistically significant differences between all groups[14].

Sub-corpus E2_SS_113 showed a higher internal variation (CV=23.64%) than sub-corpora E2_SMC_42 (CV=19.77%) and E2_SMA_38 (CV=15.79%). This is interesting considering that sub-corpus E2_SS_113 was the only sub-corpus of Discussion Essays comprising essays produced by students in one location (i.e. the UAE) compared to the other two sub-corpora containing essays produced in the UK and Mauritius. As with the other writing genres investigated thus far (5.2.1 - 5.2.3), this points to factors other than the students' geographical location that impacted the density of academic vocabulary in their Discussion Essays. Sub-corpus E2_SS_113 also contained significantly more texts (i.e. 113) than the other two sub-corpora (i.e. 42 and 38). Similar to the Exposition Essays (5.2.1), the internal variation increased with the increasing number of texts and the density of academic vocabulary decreased with an increasing internal variation.

Next, the density of academic vocabulary was explored in more depth by investigating how the density was achieved. Tables 5.14 - 5.16 show academic vocabulary which formed at least 1% of all academic tokens in all three sub-corpora of Discussion Essays.

Table 5.14: Academic vocabulary forming ≥ 1% of all academic tokens (E2_SMC_42)

Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens	Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens
*social**	42	1402	251.19	15.48	*example*	36	111	19.89	1.23
information	40	237	42.46	2.62	*encourage**	35	105	18.81	1.16
user	37	162	29.02	1.79	*report*	33	94	16.84	1.04
networking	34	152	27.23	1.68	*network*	31	93	16.66	1.03
individual	35	117	20.96	1.29	*however*	36	91	16.3	1.01

*academic vocabulary item in the assignment brief

[14] A post-hoc t-test with Bonferroni correction results: E2_SMC_42 and E2_SMA_38 (t=6.76, p<.01, d=1.47); E2_SMC_42 and E2_SS_113 (t=2.88, p<.05, d=0.52); E2_SMA_38 and E2_SS_113 (t=10.85, p<.05, d=1.96).

Table 5.15: Academic vocabulary forming ≥ 1% of all academic tokens (E2_SMA_38)

Academic lemma	Range	Raw freq.	Norm Freq.	% of ac. tokens	Academic lemma	Range	Raw freq.	Norm Freq.	% of ac. tokens
*social**	38	1264	246.13	11.6	*academic**	34	209	40.7	1.92
use	38	528	102.81	4.85	*study*	36	157	30.57	1.44
*university**	37	299	58.22	2.74	*research**	37	147	28.62	1.35
networking	35	279	54.33	2.56	*communi-cation*	36	128	24.92	1.17
information	37	222	43.23	2.04	*experience**	28	112	21.81	1.03

*academic vocabulary item in the assignment brief

Table 5.16: Academic vocabulary forming ≥ 1% of all academic tokens (E2_SS_113)

Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens	Academic lemma	Range	Raw freq.	Norm freq.	% of ac. tokens
*society**	110	1082	81.36	5.62	*technology*	81	283	21.28	1.47
information	104	526	39.55	2.73	*state*	78	226	16.99	1.17
individual	90	446	33.54	2.32	*activity*	73	211	15.87	1.1
system	80	387	29.1	2.01	*provide*	72	203	15.26	1.05
monitor	97	352	26.47	1.83	*therefore*	74	202	15.19	1.05
however	97	329	24.74	1.71	*social*	64	193	14.51	1
*benefit**	88	306	23.01	1.59					

*academic vocabulary item in the assignment brief

Tables 5.14 - 4.16 show that some of the most frequently deployed academic vocabulary items appeared in the assignment brief (including those contained in the title), again pointing to the effect of the brief and the topic on the deployment of academic vocabulary. The impact of the assignment brief was explored further by considering all academic tokens in relation to the brief (Table 5.17).

Table 5.17: Impact of assignment brief on the density of academic vocabulary (Discussion Essays)

Sub-corpus	No. of tokens	No. of ac. lemmas in brief	% of ac. lemmas from brief per ac. tokens	% of ac. lemmas from brief per all tokens
E2_SMC_42	229	41	24.13%	3.91%
E2_SMA_38	248	49	29.28%	6.21%
E2_SS_113	215	32	15.16%	2.12%

Table 5.17 shows that sub-corpus E2_SMA_38 contained the highest proportion of academic vocabulary items and all tokens formed by academic vocabulary items drawn from the assignment brief, which corresponded to the highest number of academic vocabulary items contained in the assignment brief. Conversely, sub-corpus E2_SS_113 assignment brief contained the smallest number of academic vocabulary items, corresponding to the lowest proportion of academic vocabulary formed by

academic vocabulary items drawn from the assignment brief. This points to the effect that the brief had on the deployment of academic vocabulary items in Discussion Essays (comparable to the other writing genres investigated above), whereby a higher number of academic vocabulary items contained in the brief was reflected by a higher proportion of students' texts formed by these academic vocabulary items. Hence, the academic vocabulary items drawn from the assignment briefs were explored further.

In E2_SMC_42 (Table 5.14), *social* appeared in the assignment title (*Social* media) and was used by all students with frequencies up to 53 occurrences per student (Student UK14/15/C, Figure 5.18) (M=33).

Figure 5.18: Usage of *social* Student UK14/15/C (E2_SMC_42)

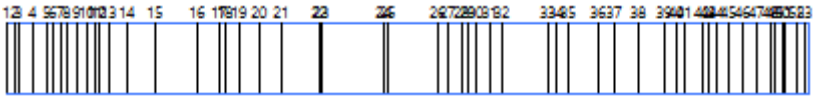


Figure 5.18 illustrates the usage of *social* throughout the essay, similar to the results from the other writing genres.

An investigation of the usage of *social* in the students' texts showed that the vast majority of instances were in combination with *media*, confirming the effect of the topic on the deployment on academic vocabulary (also found in the other writing genres investigated above). It has to be acknowledged, however, that since *social* appeared in the assignment title it was expected to be used with high frequencies in the students' texts to address the assignment topic since in the context of this assignment, *social* did not offer students many alternative expressions and was thus not expected to be replaced with synonyms.

Encourage was the second academic vocabulary items drawn from the assignment brief also appearing in the assignment question (i.e. Does social media *encourage* crime?). It is interesting to note that *encourage* was not used by all students as would be expected considering that it was contained in the assignment question. Further investigation showed that it was used with frequencies ranging from no occurrences (seven students) to eight occurrences (Student UK14/15/B) (M=3). An examination of Student UK14/15/B's essay showed that *encourage* was used to explicitly address the essay question (Figure 5.19)

Figure 5.19: Student UK14/15/B essay extract (E2_SMC_42)

...
This essay will discuss different ways that social media encourages crime and outline various situations where social media has helped to prevent crime but will conclude that social media is in fact more likely to encourage crime.
Firstly, identity theft is one of the crimes that have been encouraged by social media.
.....
This shows that social media may encourage cyber bullying as not many people witness the act unless the mentioned by the victim and the bully has a higher chance to get away with the offense.
.....
This examples suggests that social media gives hackers a platform to hack others accounts, thus it may encourage crime.
.....
Furthermore, a woman who was charged for posting a graphic video of an 8-year-old child on a social networking site is an example of how social media may encourage child pornography exploiters as it provides them with a platform to post videos. ... With the support of this type of evidence it has been seen that social media often encourages crime as it gives criminals a platform to communicate and plan riots.
.....
This essay suggests that social media has both positive and negative effects...Therefore, social media is more likely to encourage crime.

No interview or survey data were available for Student UK14/15/B to gain further insights into the students' writing strategies or perceived difficulties encountered when completing written assignments.

An examination of the seven students' essays not containing *encourage* showed that they all used alternative expression to address the essay questions (Figures 5.20 - 5.23).

Figure 5.20: Student MRU17/18/B essay extract (E2_SMC_42)

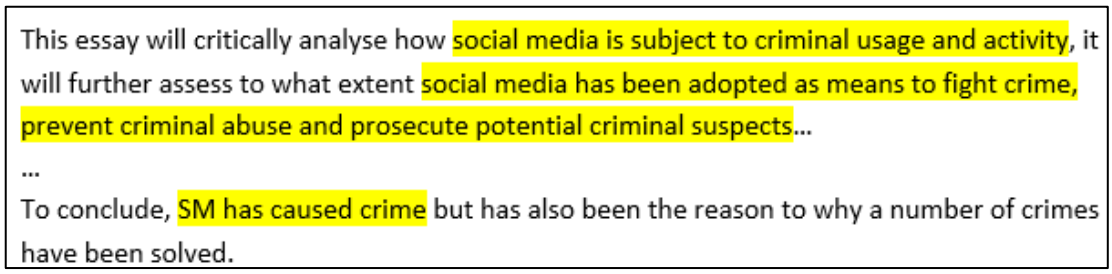
This essay will explore the relationship between social media and crime, it will analyse how social media plays a major role in influencing crime ...

To conclude, the above discussion emphasises how social media is a cornerstone for modern criminal activity.

Figure 5.21: Student MRU17/18/F essay extract (E2_SMC_42)

This essay will critically assess the ways in which social media has contributed to the growth of crime in modernity, it will further analyse how social media has become a powerful instrument to fight crime and it will strongly conclude that, given the constant rise of technology, contemporary policing of crime needs to evolve in order to tackle the challenge posed by social media.

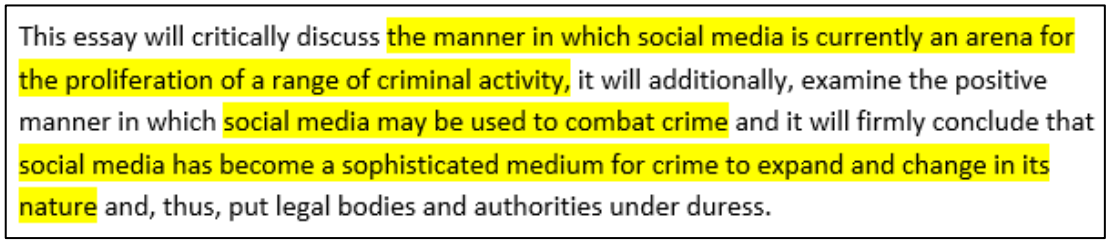
Figure 5.22: Student MRU17/18/M essay extract (E2_SMC_42)


This essay will critically analyse how social media is subject to criminal usage and activity, it will further assess to what extent social media has been adopted as means to fight crime, prevent criminal abuse and prosecute potential criminal suspects…

…

To conclude, SM has caused crime but has also been the reason to why a number of crimes have been solved.

Figure 5.23: Student MRU17/18/O essay extract (E2_SMC_42)


This essay will critically discuss the manner in which social media is currently an arena for the proliferation of a range of criminal activity, it will additionally, examine the positive manner in which social media may be used to combat crime and it will firmly conclude that social media has become a sophisticated medium for crime to expand and change in its nature and, thus, put legal bodies and authorities under duress.

Figures 5.20 - 5.23 show that the students used paraphrasing techniques instead of using *encourage*. Survey data were available for Student MRU17/8/M (Figure 5.22), which confirm that the Student used paraphrasing as one of the writing techniques when completing written assignments. Both survey and interview data were available for Student MRU17/18/O (Figure 5.23), which revealed that although the Student found paraphrasing difficult, s/he developed the strategy of using a thesaurus to paraphrase effectively during the process of completing written assignments on the IFP:

> "Paraphrasing is very challenging because this is the first time I'm hearing of it … I've never done paraphrasing in my assignments. This is the first time I'm learning how to do it, just getting each word right… Thesaurus helps a lot because you get to get a lot of different words that you can use to substitute the other words" (Student MRU17/18/O).

Similar to the findings from the other writing genres (5.2.1 - 5.2.3), this shows that students who used paraphrasing techniques tended to draw on the vocabulary contained in the assignment brief to a lesser extent.

In sub-corpus E2_SMA_38 (Table 5.15), four academic vocabulary items (*social*, *university*, *academic*, *experience)* appeared in the essay question (Could students' usage of online *social* networking for *academic* purposes have a beneficial or detrimental effect on their *experience* at *university*?). It would hence be expected that these items would be used by all students to address the assignment question.

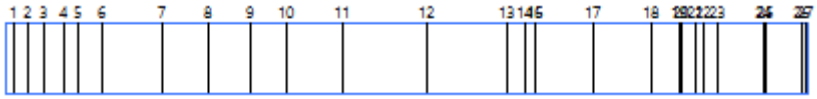
This was the case with *social* only. An investigation of *social* showed that the vast majority of instances were in combination with *media*, supporting the effect of the topic on the deployment of academic vocabulary in Discussion Essays. Similar to sub-corpus E2_SMC_42, this is unsurprising considering that *social media* was the topic under discussion in this assignment and it would thus not be expected that the students would use alternative expressions.

In terms of the second most frequently deployed academic vocabulary item contained in the assignment title and brief (*university*), only one student (Student MRU18/A) did not use this academic vocabulary item. An examination of this Student's essay revealed that s/he used paraphrasing techniques and used terms such as *tertiary education* or *higher education* instead of *university*. Survey data were available for this Student, where it was not stated that s/he encountered difficulties with using her/his own words or appropriate vocabulary, but the use of paraphrasing was selected as a writing technique employed when completing written assignments. This supports the findings from the other sub-corpora where students who used paraphrasing techniques relied on the assignment brief less than students who found using their own words and appropriate vocabulary difficult.

As for sub-corpus E2_SS_113, Table 5.16 shows that two academic vocabulary items (*society* and *benefit*) were contained in the assignment brief with *society* appearing in all four essay questions relating to this assignment from which the students could choose. These are: 1) The benefits of a surveillance society outweigh its detriments. Discuss. 2) In a surveillance society privacy is an illusion. Discuss. 3) Surveillance does more harm than good. Discuss. 4) The surveillance society is a necessity in the modern world. Discuss.

An examination of the usage of *society* showed that the majority of occurrences were in combination with *surveillance*, pointing to the impact of the assignment topic (Surveillance society) on the deployment of academic vocabulary. Further, *society* was not used by all students as would be expected considering that it was contained in the assignment question. An investigation of the usage of *society* showed that it was used with frequencies ranging from no occurrences (three students) to 27 occurrences (Student UAE17/18/C, Figure 5.24) (M=10).

Figure 5.24: Usage of *society* Student UEA17/18/C (E2_SS_113)



An examination of the three students (Students UAE17/18/II, UAE17/18/OO, UAE17/18/P) who did not use *society* showed that all three students used *surveillance* throughout the essay to refer to *surveillance society.*

The above analysis of the most frequently used academic vocabulary items appearing in the assignment brief and / or title of Discussion Essays revealed that students tended to employ two different strategies to address the assignment instruction and topic: by drawing on the academic vocabulary items contained in the assignment brief (including those contained in the assignment title), and by using paraphrasing techniques (including alternative expressions).

The following section provides a discussion of all results relating to the density of academic vocabulary across the four writing genres under investigation.

5.3 Density of academic vocabulary across writing genres: discussion

The analysis of the density of academic vocabulary in the four writing genres (5.2) has generated findings relating to the impact of the assignment brief and the assignment topic. It has also shown differences among as well as within individual writing genres. These findings are further discussed in this section.

5.3.1 The role of the assignment brief

An investigation of the density of academic vocabulary showed that some of the most frequently deployed academic vocabulary items were drawn from the assignment brief. This led to a more detailed examination of the effect of the assignment brief on the deployment of academic vocabulary items, summarised in Table 5.18 showing the number of academic lemmas contained in the assignment brief and the proportion of these academic lemmas per all academic lemmas as well as per all tokens in each sub-corpus.

Table 5.18: Impact of assignment brief on the density of academic vocabulary

Writing genre	Sub-corpus	No. of tokens	No. of ac. lemmas in brief	% of ac. lemmas from brief per ac. tokens	% of ac. lemmas from brief per all tokens
Exposition Essays	E1_LS_42	165	26	17.63%	2.95%
	E1_MI_38	156	22	16.61%	2.33%
	E1_FN_113	110	18	8.83%	1.03%
Problem Questions	PQ_BOS_48	129	15	15.7%	2.86%
	PQ_ESP_73	112	13	15.67%	2.78%
	PQ_ST_59	117	20	18.82%	3.64%
	PQ_SI_13	159	25	23.81%	3.86%
Research Reports	RR_193	455	39	29.87%	5.69%
Discussion Essays	E2_SMC_42	229	41	24.13%	3.91%
	E2_SMA_38	248	49	29.28%	6.21%
	E2_SS_113	215	32	15.16%	2.12%

Table 5.18 shows that in all writing genres, a higher number of academic vocabulary items contained in the assignment brief matched a higher proportion of academic vocabulary tokens as well as a higher proportion of students' texts being formed by these vocabulary items, confirmed by a Pearson's correlation (r=0.79 and r=0.74 respectively). This, therefore, suggests that the brief played an important role in the deployment of academic vocabulary items in the students' written assignments.

A closer examination of the most frequently deployed academic vocabulary items contained in the assignment brief showed that students tended to use two different strategies to address the assignment instructions: a) by repeatedly using the academic vocabulary items from the assignment brief; and b) by using alternative expressions (i.e. paraphrasing techniques). As for the former, an examination of the survey data showed that out of the 51 students who completed the survey, 20 stated that they used vocabulary from the assignment brief as a writing strategy when completing their written assignments. Half of these students stated that they found the use of appropriate vocabulary and / or using own words difficult. Conversely, students who used alternative ways of referring to the academic vocabulary items contained in the assignment title did not state they found the use of appropriate vocabulary or own words difficult.

This, to some extent, suggests that drawing on vocabulary from the assignment brief may have been a strategy used by students who were less confident in using appropriate vocabulary and own words; that is, less proficient IFP student writers tended to draw on the vocabulary contained in the brief to a greater extent. This aligns with Milton's (2001) findings relating to the reliance on assignments prompts

by Hong Kong Chinese university students. His study showed that the least proficient Hong Kong students relied particularly heavily on the vocabulary items contained in the assignment prompts (2.2.3 & 3.2.2).

The reliance of IFP students on the vocabulary items contained in the assignment brief may thus have been a compensatory strategy of novice writers in English, highlighting the important role that the assignment brief plays in the students' written production. However, it also ought to be pointed out that the students needed to clearly demonstrate that they were following the instructions set out in the assignment brief in order to meet the assignments' requirements, and integrating academic vocabulary contained in the assignment brief achieved this. Thus, drawing on the academic vocabulary from the assignment brief needed not necessarily be a strategy deployed by less proficient writers only.

The importance of the assignment brief was also emphasised by Gustafson-Pearce (2009, p. 2585) who regards the brief as "a vital component for modules, which set assignments as a method of assessment"… and "[a]s such, it becomes one of the major tools of the module learning aids". This is corroborated by Russell (2001, p. 261) who notes that one of the factors shaping higher education students' writing are the pedagogical tools with which the students are provided. The findings of this study thus suggest that the assignment brief was used as one of the pedagogical tools on which students drew in the process of completing their written assignments as it was utilised as a source of academic vocabulary.

The use of the assignment brief as a pedagogical tool, whereby it served as a source of vocabulary in this study, can be seen as somewhat related to findings generated by research into writing from sources (e.g., textual borrowing, source-based writing, reading-to-write constructs or integrated writing tasks) (3.2.2). This research shows that one of the functions that sources serve is "as a language repository" (Plakans & Gebril, 2012, p. 18) and conversely, that one of the disadvantages of writing without sources is lack of topic-related vocabulary (Leki & Carson, 1997). Therefore, the role of the assignment brief ought not to be underestimated or overlooked in the context of assessed academic writing as this study has shown that the function of an assignment brief was not only to provide instructions and requirements of a writing task, but it also served the less obvious and perhaps unintended function of providing students with appropriate vocabulary to be integrated in their writing.

5.3.2 The role of the topic

The assignment topic was found to be another factor impacting the density of academic vocabulary across all four writing genres with some of the most frequently deployed academic vocabulary items contained in the assignment title. The impact of the topic on vocabulary deployment is in accord with Olinghouse and Wilson's (2013) findings showing that students consider the topic when selecting vocabulary to integrate in their writing as the topic knowledge is one of the "knowledge bases students access to select appropriate words while composing written text" (p. 59) (3.2.5).

As noted previously, however, it has to be acknowledged that some of the vocabulary items contained in the assignment title not only offered students few alternatives, but it can also be assumed that these vocabulary items had to be used by students in addressing the assignment topics and questions. This was likely one of the factors resulting in rather high density of these academic vocabulary items drawn from the assignment titles. The findings of this study, therefore, underline the extent of the effect that the topic had on the density of academic vocabulary in the students' written assignments with the academic vocabulary items drawn from the assignment titles being among those occurring with the highest density.

These academic vocabulary items relating to the topic reflect Nation's (2001) category of 'technical words' (3.2.1) referring to "words that are very closely related to the topic and subject area of the text" and "differ from subject area to subject area" (Nation, 2001, p. 12). The prevalence of the identified academic vocabulary closely related to the assignment topics thus suggests that within the category of general academic vocabulary (i.e. words distinctive to academic language common across a range of disciplines) exist technical words; that is, in this study academic vocabulary also subsumes technical words unique to different topics. Hence, this finding does not reflect the distinction between technical and academic vocabulary where "academic vocabulary is comprised of words used in a variety of academic contexts … rather than words associated with specific disciplines or topics" (Olinghouse & Wilson, 2013, p. 47), but instead points to an intersection between these two vocabulary types.

This became evident from the titles under Assignment 2 (the Problem Question assignments), for instance, which belonged to the same writing genre, but their

topics would typically be used in different subject areas or disciplines (e.g. psychology or business) reflecting this writing genre's social function of preparing for professional practice (Table 2.1 in 2.2.4). Consequently, some of the identified general academic vocabulary items displayed different meanings across these topic and subject areas. For instance, *control* occurring in the Ethics in Social Psychology assignment (PQ_ESP_73) as well as in the Business Organisational Structure assignment (PQ_BOS_48) referred to a *control group* in PQ_ESP_73, whereas in PQ_BOS_48 *control* referred to a *span of control*. This example clearly demonstrates the difference in meaning that general academic vocabulary can carry across different subject areas as it illustrates that although belonging to the same writing genre, academic vocabulary displays different meanings reflecting the assignment's topic which is closely related to a specific subject area / discipline.

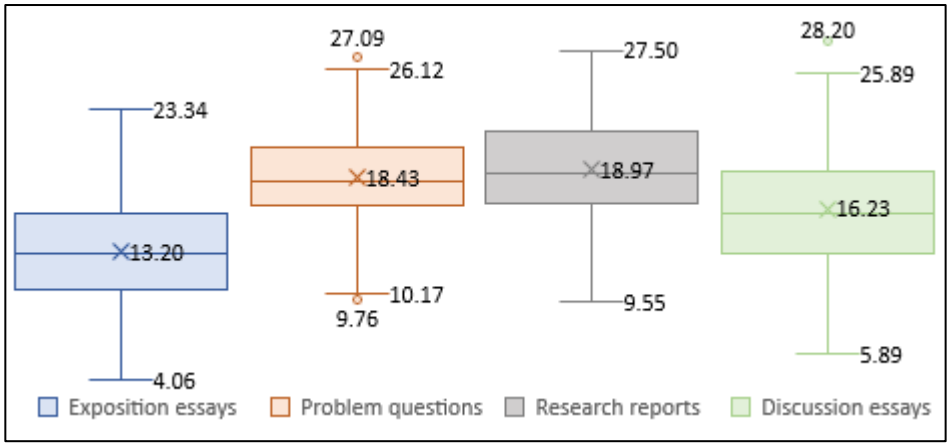
These differences may indicate that the label of 'general academic vocabulary', referring to words common across a wide range of academic disciplines, may be misleading as the topic and / or subject area seem to play an important role in the behaviour and meaning of these vocabulary items. This is supported by several corpus-based studies investigating general academic vocabulary (e.g., Durrant, 2014, 2016; Hyland & Tse, 2007, 2009; Martínez et al., 2009) which argue that "there may not be a usefully large set of vocabulary which is frequent across disciplines" (Durrant, 2016, p. 50). This argument often subsumes two important points: one relating to general academic vocabulary items not being evenly distributed across disciplines and the other concerning the different uses of general academic vocabulary across disciplines (Durrant, 2014, p. 329) (3.2.4). This study's findings corroborate both points: the academic vocabulary items in the different sub-corpora of Problem Questions (which reflect various subject areas / disciplines) showed the usage of different academic vocabulary items as well as different frequencies with which these different academic vocabulary items were used; and the different usage, behaviour and meaning of academic vocabulary across fields of study was demonstrated by the deployment of *control,* which carried different meanings in different disciplines reflecting the subject-specific topical area. This point accords with Hyland and Tse's (2007) finding that words "often occur and behave in different ways across disciplines in terms of range, frequency, collocation, and meaning" (p. 235).

The findings relating to the role of the assignment topic hence showed that the effect of the topic on the density of academic vocabulary in the students' written assignments appeared to be greater than the impact of the writing genre itself. This could be seen from the varied density of academic vocabulary that existed within the different sub-corpora belonging to the same writing genre (Figures 5.2, 5.10 and 5.17 in 5.2) as well as from the deployment of the different academic vocabulary items displaying the highest density of academic vocabulary (presented in 5.2). Further, it was found that among the most frequently deployed academic vocabulary items were so-called technical words, contrary to the distinction between academic and technical words. These findings are in agreement with Hyland and Tse's (2007, 2009) argument relating to semantic variations of words resulting from the disciplinary context in which they are used. These findings thus underline the importance of knowledge of topic-related vocabulary items, their behaviour and meaning in the specific contexts in which they are used. In the context of generic IFPs, however, it is important to find a balance between focusing on topic-specific vocabulary necessary for completion of a specific task (often related to a specific discipline, which may not be of equal usefulness to all learners) and general academic vocabulary common in various academic contexts.

5.3.3 The role of the genre

As shown in Figure 5.1 (in section 5.2) and replicated in Figure 5.25 below, differences existed in the density of academic vocabulary among the four writing genres investigated in this study with the Research Reports displaying the highest and the Exposition Essays the lowest density of academic vocabulary.

Figure 5.25: Density of academic vocabulary (summary overview)



The boxplot in Figure 5.25 may thus suggest that the writing genre impacted the density of academic vocabulary. However, an analysis of individual sub-corpora by assignment titles within these writing genres (sections 5.2.1, 5.2.2 and 5.2.4) showed significant differences in the density of academic vocabulary within these writing genres (with the exception of the writing genre of Research Reports which did not comprise sub-corpora), summarised in Figure 5.26.

Figure 5.26: Density of academic vocabulary (detailed overview)

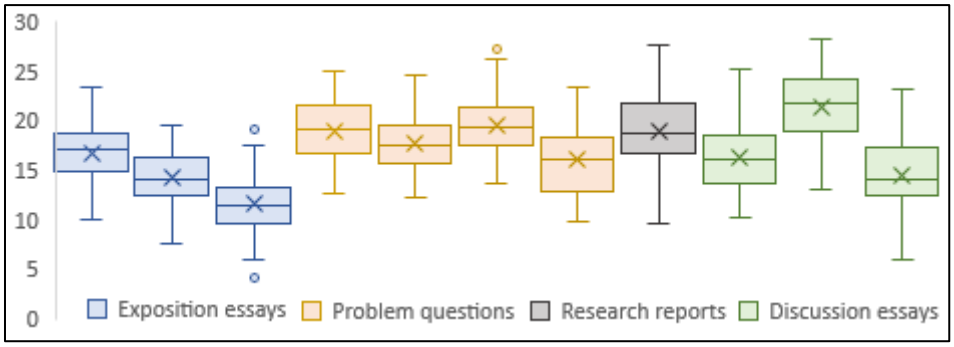


Figure 5.26 thus points to factors other than the writing genre that impacted on the density of academic vocabulary in the students' assessed academic writing. From the above-discussed findings, the assignment brief (5.3.1) and topic (5.3.2) seemed to have a greater effect on the density of academic vocabulary than the writing genre. This is in contradiction with Csomay and Prades's (2018, p. 114) findings showing that "text type and rhetorical purpose play a greater role in the percentage of academic words than merely topic selection" meaning that "the genre that students are asked to produce influences academic word use more than topic" (3.2.5).

A possible reason why this study's finding does not corroborate Csomay and Prades's (2018) may be the fact that this study was able to investigate not only different writing genres, but also various topics within the same writing genre, which had the potential to show how the different topics impacted on the density of academic vocabulary within a writing genre. This highlights an important point relating to an investigation of vocabulary deployment in student writing produced in response to one topic only, in that this may not be a sufficient base for an exploration of the impact of the writing genre versus topic on the deployment of academic vocabulary.

Thus, the role that the writing genre plays in the density of academic vocabulary in the students' written assignments seemed to be less significant than the assignment topic and brief. In addition, considerable differences in internal variation among sub-corpora of the same writing genre were found and it was shown that the density of

academic vocabulary decreased with the increasing internal variation. Hence, internal variation seemed to be a further factor that impacted on the overall density of academic vocabulary to a greater extent than the writing genre. It is also interesting to note that in the Exposition and Discussion Essays, the internal variation increased with a higher number of texts, pointing to a relationship between the number of texts and internal variation, which in turn impacted on the density of academic vocabulary in essays.

In sum, it can be said that the writing genre did impact on the deployment of academic vocabulary but to a lesser degree than some other factors including the assignment topic, assignment brief, the internal variation and number of texts; the assignment topic and brief seemed to have a direct impact on the usage of academic vocabulary by students, whereas the number of texts seemed to impact on internal variation, which in turn had an effect on the overall density of academic vocabulary in individual sub-corpora.

5.3.4 Density of academic vocabulary across writing genres: summary

The exploration of the usage of academic vocabulary from the perspective of density of academic vocabulary across the four writing genres of the IFP students' assessed academic writing showed the impact of the academic vocabulary items contained in the assignment brief, on which students tended to draw when completing their written assignments. In particular, students less confident in using appropriate vocabulary, their own words and paraphrasing tended to use the assignment brief as a vocabulary repository. The assignment topic was found to be a further factor that impacted on the density of academic vocabulary, whereby a large proportion of academic vocabulary was formed by academic vocabulary items contained in the assignment title. These academic vocabulary items can also be regarded as technical vocabulary items, which in this study were subsumed in academic vocabulary. The writing genre also was shown to play a role in the density of academic vocabulary with certain writing genres displaying a higher density overall; however, this impact was found to be less significant than the role of the assignment brief and topic.

5.4 Diversity of academic vocabulary: results

This section provides the results relating to the diversity of academic vocabulary in the four writing genres under investigation. First, an overview of the diversity in the four writing genres is provided (Figure 5.27), followed by an analysis of the diversity of academic vocabulary in individual sub-corpora forming the different writing genres.

Figure 5.27: Diversity of academic vocabulary across writing genres

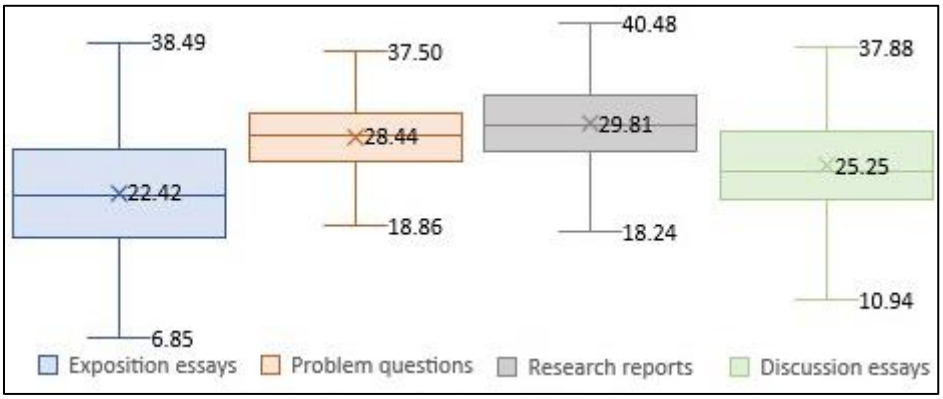


Figure 5.27 displays the mean average of the diversity of academic vocabulary in the four writing genres (i.e. 22.42% in Exposition Essays, 28.44% in Problem Questions, 29.81% in Research Reports, 25.25% in Discussion Essays) and the range showing the overall dispersion of the values in each writing genre. Unlike the density of academic vocabulary (Figure 5.1 in 5.2), no outliers were identified.

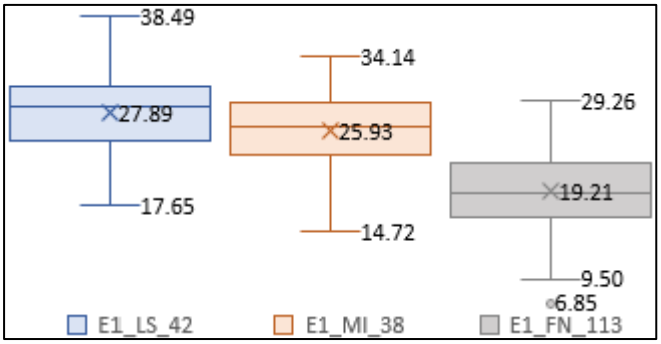
Similar to the density of academic vocabulary (Figure 5.1 in 5.2), the writing genre of Exposition Essays displayed the lowest and the Research Reports the highest diversity of academic vocabulary (Figure 5.27). A one-way ANOVA showed statistically significant differences between the writing genres ($f=88.46$, $p<.05$) with a post-hoc t-test with Bonferroni correction confirming statistically significant differences between all pairs of the four sub-corpora with a small effect size, however, between the Problem Questions and Research Reports[15]. The individual writing genres are further explored next.

[15] A post-hoc t-test with Bonferroni correction results: Exposition Essays and Problem Questions ($t=12.08$, $p<0.01$, $d=1.19$), Exposition Essays and Research Reports ($t=14.83$, $p<0.01$, $d=1.41$), Exposition Essays and Discussion Essays ($t=5.67$, $p<0.01$, $d=0.51$), Problem Questions and Research Reports ($t=2.75$, $p<0.05$, $d=0.33$), Problem Questions and Discussion Essays ($t=6.41$, $p<0.01$ $d=0.71$), Research Reports and Discussion Essays ($t=9.16$, $p<0.01$, $d=0.97$).

5.4.1 Assignment 1 Exposition Essays

This section presents results relating to the writing genre of Exposition Essays comprising the following three sub-corpora: E1_LS_42 (Learning Styles); E1_MI_38 (Multiple Intelligences); and E1_FN_113 (Fake News) (Figure 5.28).

Figure 5.28: Diversity of academic vocabulary (Exposition Essays)



E1_LS_42: SD=4.76%, CV=17.05%, 95% CI=[26.41, 29.37]
E1_MI_38: SD=4.58%, CV=17.68%, 95% CI=[24.42, 27.44]
E1_FN_113: SD=4.53, CV=23.59%, 95% CI=[18.37, 20.05]

Figure 5.28 shows that the diversity of academic vocabulary in the three sub-corpora of Exposition Essays ranged from 6.85% (E1_FN_113) to 38.49% (E1_LS_42), with sub-corpus E1_FN_113 showing the lowest (M=19.21%) and E1_LS_42 showing the highest (M=27.89%) diversity of academic vocabulary. A one-way ANOVA confirmed that the differences between the three sub-corpora were statistically significant (f=68.61, p<.00001) and a post-hoc t-test with Bonferroni correction confirmed statistically significant differences between several sub-corpora[16].

Sub-corpus E1_FN_113 displayed a higher internal variation (CV=23.59%) than sub-corpora E1_LS_42 (CV=17.05%) and E1_MI_38 (CV=17.68%). Similar to the density of academic vocabulary (5.2), this is somewhat surprising considering that sub-corpus E1_FN_113 was the only sub-corpus of Exposition Essays containing essays produced by students in one location (UAE). Thus, in line with the findings relating to the density of academic vocabulary (5.2), this indicates that factors other than the students' place of study impacted the diversity of academic vocabulary in their written production. Hence, the higher number of texts contained in sub-corpus E1_FN_113 could have been a factor impacting internal variation. The above thus

[16] A post-hoc t-test with Bonferroni correction results: statistical significant differences found between E1_LS_42 and E1_FN_113 (t=10.47, p<0.1, d=1.87), and between E1_MI_38 and E1_FN_113 (t=7.81, p<.01, d=1.48) but not between E1_LS_42 and E1_MI_38 (t=1.91, p≥0.5, d=0.42).

points to a relationship between a higher internal variation and a lower diversity of academic vocabulary (i.e. the diversity of academic vocabulary decreased with an increasing internal variation), and between the number of texts and internal variation (i.e. internal variation increased with the increasing number of texts).

The next section explores the diversity of academic vocabulary in more detail by investigating how the diversity was achieved. Table 5.19 shows how the diversity of academic vocabulary was distributed in relation to student numbers; in other words, what percentage of academic vocabulary types per all academic types was generated by how many students.

Table 5.19: Academic vocabulary types distribution (Exposition Essays)

% of students	E1_LS_42		E1_MI_38		E1_FN_113	
	No of ac. types	(% of ac. types)	No of ac. types	(% of ac. types)	No of ac. types	(% of ac. types)
≥75%	7	(0.9%)	5	(0.71%)	3	(0.34%)
≥50%	30	(3.86%)	32	(4.56%)	20	(2.29%)
≥25%	96	(12.36%)	108	(15.41%)	40	(4.58%)
<25%	681	(87.64%)	593	(84.59%)	834	(95.42%)
Individual SS	306	(39.38%)	258	(36.8%)	317	(36.27%)

Table 5.19 shows that in all three sub-corpora, the diversity of academic vocabulary was achieved by a small number (<25%) of students; that is, less than 25% of students generated between approximately 85% - 95% of academic vocabulary types and less than 1% of academic vocabulary types were shared by more than 75% of students. This indicates that the diversity of academic vocabulary in Exposition Essays was, to a large extent, generated by a small number of students. This suggests that there was only a very small set of academic vocabulary types shared by the students (further discussed in 5.5.1). The academic vocabulary types shared by the majority of the students (i.e. at least 50% of students) in individual sub-corpora of Exposition Essays are shown in Tables 5.20 - 5.22.

Table 5.20: Academic vocabulary used by at least 50% of students (E1_LS_42)

Academic lemma	Range	Raw freq.	Norm freq.	Function	Academic lemma	Range	Raw freq.	Norm freq.	Function
*experience**	42	259	76.65	RO-top	*relate**	24	41	12.13	RO-proc
*theory**	36	164	48.54	RO-top	*strategy*	23	67	19.83	RO-top
study	35	117	34.63	RO-top	*active*	23	50	14.8	RO-top
information	33	89	26.34	RO-top	*task*	22	55	16.28	RO-top
process	33	69	20.42	RO-proc	*approach*	22	39	11.54	RO-top
knowledge	33	60	17.76	RO-top	*discussion*	22	38	11.25	RO-proc
group	32	114	33.74	RO-desc	*observation*	22	38	11.25	RO-top
*type**	29	72	21.31	RO-desc	*method*	22	36	10.65	RO-top
improve	28	53	15.69	R0-proc	*conclude*	22	29	8.58	TO-struc
individual	27	50	14.8	RO-desc	*furthermore*	22	25	7.4	TO-trans
*identify**	26	58	17.17	RO-proc	*characteristic**	21	43	12.73	RO-desc
subject	25	58	17.17	RO-top	*describe*	21	35	10.36	RO-proc
*weakness**	25	48	14.21	RO-top	*reflect*	21	35	10.36	RO-proc
*example**	24	51	15.09	TO-struc	*base*	21	34	10.06	RO-desc
experiment-ation	24	51	15.09	RO-top	*conclusion*	21	27	7.99	TO-struc

*academic vocabulary item in the assignment brief

Table 5.21: Academic vocabulary used by at least 50% of students (E1_MI_38)

Academic lemma	Range	Raw freq.	Norm freq.	Function	Academic lemma	Range	Raw freq.	Norm freq.	Function
*theory**	36	147	41.61	RO-top	*spatial*	23	44	12.46	RO-top
*experience**	35	91	25.76	RO-top	*identify**	23	43	12.17	RO-proc
*multiple**	34	113	31.99	RO-top	*type*	22	72	20.38	RO-desc
logical	28	71	20.1	RO-top	*relate**	22	42	11.89	RO-proc
*example**	28	68	19.25	TO-struc	*process*	22	39	11.04	RO-proc
information	27	50	14.15	RO-top	*state*	22	39	11.04	TO-struc
*university**	27	50	14.15	RO-loc	*practice**	22	36	10.19	RO-proc
subject	26	66	18.68	RO-top	*conclude*	22	24	6.79	TO-struc
develop	25	75	21.23	RO-proc	*reflect**	21	41	11.61	RO-proc
knowledge	25	51	14.44	RO-top	*discuss**	21	28	7.93	TO-struc
individual	24	68	19.25	RO-desc	*conclusion*	21	21	5.94	TO-struc
group	24	60	16.98	RO-desc	*value*	20	24	6.79	RO-proc
apply	24	53	15	RO-proc	*define*	19	42	11.89	TO-fram
improve	23	73	20.66	RO-proc	*human*	19	29	8.21	RO-desc
possess	23	59	16.7	RO-desc	*perform*	19	28	7.93	RO-proc
therefore	23	47	13.3	TO-res	*result*	19	27	7.64	TO-res

*academic vocabulary item in the assignment brief

Table 5.22: Academic vocabulary used by at least 50% of students (E1_FN_113)

Academic lemma	Range	Raw freq.	Norm freq.	Function	Academic lemma	Range	Raw freq.	Norm freq.	Function
*harmful**	109	410	53.13	RO-top	*source*	64	129	16.72	RO-top
social	101	414	53.65	RO-top	*society*	63	187	24.23	RO-desc
information	98	360	46.65	RO-top	*firstly*	62	64	8.29	TO-struc
affect	78	211	27.34	RO-proc	*article*	60	194	25.14	RO-top
*discuss**	76	105	13.61	TO-struc	*therefore*	60	109	14.13	TO-res
*conclusion**	75	77	9.98	TO-struc	*influence*	59	107	13.87	RO-proc
effect	71	133	17.24	RO-top	*report*	59	98	12.7	RO-top
state	70	165	21.38	TO-struc	*impact*	58	122	15.81	RO-top
example	68	113	14.64	TO-struc	*result*	56	128	16.59	TO-res
furthermore	65	71	9.2	TO-trans	*moreover*	56	69	8.94	TO-trans

*academic vocabulary item in the assignment brief

In total, Tables 5.20 – 5.22 contain 82 academic vocabulary items: Table 5.20 shows 30 academic types, out of which 8 (27%) were contained in the assignment brief. Table 5.21 contains 32 academic vocabulary items with 10 (31%) appearing in the assignment brief. In Table 5.22, 20 academic vocabulary types were used by at least 50% of the students, out of which 3 (15%) appeared in the assignment brief. Hence, the academic vocabulary types among those used by the majority of the students were formed to some extent (15% - 31%) by academic vocabulary contained in the assignment brief (further discussed in 5.5.1).

However, as can be seen from Tables 5.20 - 5.22, some of the academic vocabulary items deployed by the majority of students were used in more than one sub-corpus of Exposition Essays (e.g. *conclude*, *example*, *information*). For the purpose of obtaining an accurate number of the core academic vocabulary, these duplicated vocabulary items were counted once only, resulting in a core of 58 academic vocabulary items deployed across the three titles of Exposition Essays (with 24 duplicated items removed).

An analysis of the meaning or function of the academic vocabulary (based on Hyland's classification presented in 4.4.2) deployed by at least 50% of students (58 academic vocabulary items in all three sub-corpora) showed that approximately 41% of these vocabulary items were research-oriented vocabulary relating to the assignment topic (Table 5.23).

Table 5.23: Academic vocabulary function (Exposition Essays)

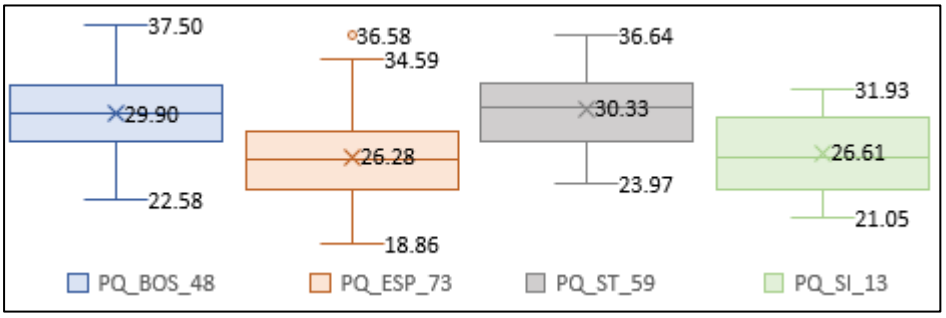
Function	No (%)	Function	No (%)
RO-top	24 (41.38%)	TO-res	2 (3.45%)
RO-proc	14 (24.14%)	TO-trans	2 (3.45%)
RO-desc	8 (13.79%)	RO-loc	1 (1.72%
TO-struc	6 (10.34%)	TO-fram	1 (1.72%)

The above analysis hence points to the effect that both the topic and assignment brief had on the diversity of academic vocabulary in IFP students' Exposition Essays (further discussed in 5.5.1 and 5.5.2).

5.4.2 Assignment 2 Problem Questions

This section presents results relating to the writing genre of Problem Questions comprising four sub-corpora: PQ_BOS_48 (Business Organisational Structure); PQ_ESP_73 (Ethics in Social Psychology); PQ_ST_59 (Sustainable Tourism); and PQ_SI_13 (Sustainable Initiatives) (Figure 5.29).

Figure 5.29: Diversity of academic vocabulary (Problem Questions)



PQ_BOS_48: SD=3.5, CV=11.71%, 95 CI =[28.88, 30.92]
PQ_ESP_73: SD=3.74, CV=14.22%, 95% CI=[25.41, 27.15]
PQ_ST_59: SD=3.04, CV=10%, 95% CI=[29.54, 31.12]
PQ_SI_13: SD=3.56, CV=13.37%, 95% CI=[24.46, 28.76]

Figure 5.29 shows that the diversity of academic vocabulary in the four sub-corpora of Problem Questions ranged between 18.86% (PQ_ESP_73) and 37.50% (PQ_BOS_48), with sub-corpus PQ_ESP_73 showing the lowest (M=26.28%) and PQ_ST_59 showing the highest (M=30.33%) diversity of academic vocabulary. A one-way ANOVA showed that the differences between the four sub-corpora were statistically significant (f=19.41, p<.00001) and a post-hoc t-test with Bonferroni correction showed statistically significant differences between several sub-corpora[17]:

[17] Statistically significant differences with large effect sizes found between the following sub-corpora: PQ_BOS_48 and PQ_ESP_73 (t=5.63, p<.01, d=0.99), PQ_BOS_48 and PQ_SI_13 (t=3.04, p<.05, d=0.94), PQ_ESP_73 and PQ_ST_59 (t=6.69, p<.01, d=1.18), and PQ_ST_59 and PQ_SI_13 (t=3.51, p<.01, d=1.19).

Sub-corpus PQ_ESP_73 showed the highest internal variation (CV=14.22%). This sub-corpus contained the highest number of texts (i.e. 73) and also displayed the lowest diversity of academic vocabulary. In addition, unlike sub-corpora PQ_BOS_48 and PQ_ST_59 (which showed the lowest internal variation and contained texts produced by students in two geographical locations), sub-corpus PQ_ESP_73 contained texts produced by students in one place of study (UAE). Hence, this suggests that factors other than the students' geographical location impacted the diversity of academic vocabulary in the students' written production. Similar to the sub-corpus of Exposition Essays, the diversity of academic vocabulary decreased with the increasing internal variation.

The next section explores the diversity of academic vocabulary by investigating how the diversity was achieved. Table 5.24 shows how the diversity of academic vocabulary was distributed in relation to student numbers.

Table 5.24: Academic vocabulary types distribution (Problem Questions)

% of students	PQ_BOS_48		PQ_ESP_73		PQ_ST_59		PQ_SI_13	
	No of ac. types	(% of ac. types)	No of ac. types	(% of ac. types)	No of ac. types	(% of ac. types)	No of ac. types	(% of ac. types)
≥75%	25	(2.43%)	26	(2.49%)	48	(4.2%)	30	(5.67%)
≥50%	73	(7.1%)	55	(5.27%)	95	(8.32%)	81	(15.31%)
≥25%	173	(16.83%)	138	(13.23%)	220	(19.26%)	187	(35.35%)
<25%	855	(83.17%)	905	(86.77%)	922	(80.74%)	342	(64.65%)
Individual SS	337	(32.78%)	347	(33.27%)	353	(30.91%)	260	(49.15%)

Table 5.24 shows that the diversity of academic vocabulary was generated predominantly by a small number (<25%) of students: less than 6% of academic vocabulary types were shared by more than 75% of students. In three of the four sub-corpora, less than 8.5% of academic vocabulary types were shared by more than 50% of students, while in one sub-corpus (PQ_SI_13) this was nearly twice as many. This is interesting as sub-corpus PQ_SI_13 contained a significantly smaller number of texts compared to the other sub-corpora. Similarly, in the same three sub-corpora, less than 19.5% of academic vocabulary types were generated by 25% of students, while nearly twice as many academic vocabulary types (over 35%) were generated by 25% of students in sub-corpus PQ_SI_13. Sub-corpus PQ_SI_13 also differed in the number of academic vocabulary types generated by less than 25% of students, where a significantly lower number of academic vocabulary types was generated by those students (around 65% of academic vocabulary types compared to 80% - 87% in the

other three sub-corpora). Sub-corpus PQ_SI_13 also displayed the highest proportion of academic vocabulary types generated by individual students (nearly 50% compared to around 30% in the other three sub-corpora). However, caution is needed when interpreting the results relating to sub-corpus PQ_SI_13 due to the small number of texts it contained in comparison to the other three sub-corpora of Problem Questions.

Overall, the above indicates that the diversity of academic vocabulary in Problem Questions was, to a large extent, generated by a small number of students, suggesting that there was only a very small set of academic vocabulary types shared by the students in this writing genre, further discussed in 5.5.1.

The academic vocabulary types shared by the majority of the students (i.e. at least 50% of students) in the Problem Question assignments are included in Appendices 7a – 7d, totalling 304 academic vocabulary items across the four sub-corpora of Problem Questions: The Tables in Appendices 7a - 7d contain 73 (Appendix 7a), 55 (Appendix 7b), 95 (Appendix 7c) and 81 (Appendix 7d) academic vocabulary items, out of which 8 (11%), 7 (13%), 13 (14%) and 14 (17%) items (respectively), were contained in the assignment brief. Thus, the academic vocabulary types among those used by the majority of the students were formed to some extent (11% - 17%) by academic vocabulary contained in the assignment brief. This points to the effect of the brief on the deployment of these academic vocabulary items, further discussed in 5.5.1.

Out of the 304 academic vocabulary items identified across the four sub-corpora, 149 were found to have been deployed in more than one sub-corpus (e.g. *aim*, *environment*, *sustainability*). These overlapping items were counted once only, resulting in a core of 155 academic vocabulary used in the Problem Question genre.
An analysis of the function of the core 155 academic vocabulary items showed that around one third were research-oriented procedural vocabulary items. The second most frequent function was research-oriented descriptive vocabulary (around 31%) and third most frequent was research-oriented vocabulary relating to the assignment topic (approximately 18%) (see Table 5.25 for remaining functions). This is significantly less than the research-oriented topical academic vocabulary identified in the sub-corpus of Exposition Essays (i.e. 41%, Table 5.23). This is likely due to the different social purposes of the writing genres and their other characteristics (e.g. their structure / stages) (2.2.4), further discussed in 5.5.2.

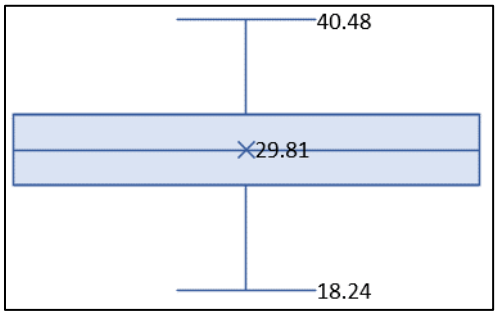
Table 5.25: Academic vocabulary function (Problem Questions)

Function	No. (%)	Function	No. (%)
RO-proc	52 (33.55%)	TO-fram	5 (3.23%)
RO-desc	48 (30.97%)	TO-res	4 (2.58%)
RO-top	28 (18.06%)	TO-trans	3 (1.94%)
TO-struc	14 (9.03%)	RO-loc	1 (0.65%)

5.4.3. Assignment 3 Research Report

This section presents results relating to the writing genre of Research Reports focusing on the topic of 'Managing barriers to sustainability' (Figure 5.30).

Figure 5.30: Diversity of academic vocabulary (Research Reports)



SD=4.36, CV=14.63%, 95% CI=[29.19, 30.43]

Figure 5.30 shows that the diversity of academic vocabulary ranged between 18.24% - 40.48% with an average diversity of 29.81% and some internal variation (CV=14.63%). Similar to the density of academic vocabulary in this sub-corpus (5.2.3), since this was the only sub-corpus containing texts from all students across all three campuses, it is interesting that the internal variation was lower than in some sub-corpora containing fewer texts and / or texts from more than one campus (e.g. all Exposition Essays, 5.4.1). This suggests that factors other than the students' place of study impacted the diversity of academic vocabulary in this writing genre.

Next, Table 5.26 shows how the diversity of academic vocabulary was distributed.

Table 5.26: Academic vocabulary types distribution (Research Reports)

% of students	No of ac. types (% of ac. types)	
≥75%	18	(1.14%)
≥50%	66	(4.18%)
≥25%	167	(10.58%)
<25%	1412	(89.42%)
Individual SS	394	(24.95%)

Table 5.26 shows that only a small proportion of academic vocabulary types (i.e. approximately 4%) was shared by the majority (i.e. at least 50%) of students, while

less than 25% of students generate nearly 90% of academic vocabulary types. The academic vocabulary types shared by at least 50% of students included 66 academic vocabulary types, out of which 24 (i.e. 36%) were drawn from the assignment brief (Appendix 7e).

An analysis of the function of the academic vocabulary used by the majority of students (Table 5.27 below) shows that one third were research-oriented procedural academic vocabulary items, whereas academic vocabulary relating to the assignment topic formed less than 8%. The reason for a smaller number of research-oriented topical vocabulary is likely to be the fact that within the general topic of barriers to sustainability, the students chose to focus on different barriers (4.3.4) generating fewer topic specific academic vocabulary. Further, the social purposes of this writing genre (2.2.4) were likely to have influenced the different functions of these academic vocabulary items, further discussed in 5.5.2.

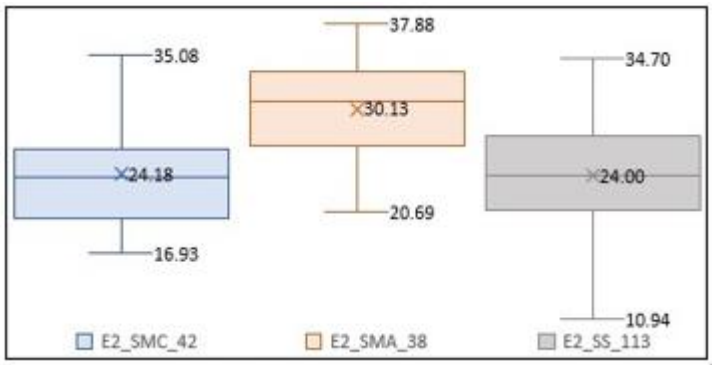
Table 5.27: Academic vocabulary function (Research Reports)

Function	No (%)	Function	No (%)
RO-proc	22 (33.33%)	TO-fram	2 (3.03%)
RO-desc	19 (28.79%)	TO-trans	2 (3.03%)
TO-struc	15 (22.73%)	TO-res	1 (1.52%)
RO-top	5 (7.58%)		

5.4.4 Assignment 4 Discussion Essays

This section presents the results relating to the writing genre of Discussion Essays comprising the following three sub-corpora: E2_SMC_42 (Social Media and Crime); E2_SMA_38 (Social Media for Academic Purposes); and E2_SS_113 (Surveillance Society) (Figure 5.31).

Figure 5.31: Diversity of academic vocabulary (Discussion Essays)



E2_SMC_42: SD=4.32, CV=17.84%, 95% CI=[22.84, 25.52]
E2_SMA_38: SD=4.2, CV=13.94%, 95% CI= [28.75, 31.51]
E2_SS_113: SD=4.57, CV=19.06%, 95% CI=[23.15, 24.85]

Figure 5.31 shows that the diversity of academic vocabulary in the three sub-corpora of Discussion Essays ranged between 10.94% (E2_SS_113) and 37.88% (E2_SMA_38), with sub-corpus E2_SS_113 showing the lowest (M=24%) and E2_SMA_38 the highest (M=30.13%) diversity of academic vocabulary. A one-way ANOVA confirmed that the differences between the three sub-corpora were statistically significant (f= 28.54, p<.001) and a post-hoc t-test with Bonferroni correction showed statistically significant differences between several sub-corpora[18].

Sub-corpus E2_SS_113 showed a higher internal variation (CV=19.06%) than sub-corpora E2_SMC_42 (CV=17.84%) and E2_SMA_38 (CV=13.94%). This is somewhat surprising considering that sub-corpus E2_SS_113 was the only sub-corpus of Exposition Essays containing essays produced by students in one location (UAE). Similar to the density of academic vocabulary in this sub-corpus (5.2.1), internal variation decreased with the decreasing number of texts. Thus, the higher number of texts contained in sub-corpus E2_SS_113 could have been a factor impacting internal variation. The above hence points to a relationship between a higher internal variation and a lower diversity of academic vocabulary (i.e. the diversity of academic vocabulary decreased with the increasing internal variation), and between the number of texts and internal variation (i.e. internal variation increased with the increasing number of texts).

The next section investigates how the diversity of academic vocabulary was achieved. Table 5.28 displays how the diversity of academic vocabulary was distributed in relation to student numbers.

Table 5.28: Academic vocabulary types distribution (Discussion Essays)

% of students	E2_SMC_42		E2_SMA_38		E2_SS_113	
	No of ac. types (% of ac. types)		No of ac. types (% of ac. types)		No of ac. types (% of ac. types)	
≥75%	13	(1.48%)	26	(3.06%)	7	(0.59%)
≥50%	45	(5.12%)	70	(8.23%)	29	(2.45%)
≥25%	131	(14.9%)	186	(21.86%)	104	(8.79%)
<25%	748	(85.1%)	665	(78.14%)	1079	(91.21%)
Individual SS	312	(35.49%)	295	(34.67%)	354	(29.92%)

Table 5.28 shows that in all three sub-corpora, the diversity of academic vocabulary was generated by a small number (<25%) of students who produced between

[18] A post-hoc t-test with Bonferroni correction showed a statistically significant differences between E2_SMC_42 and E2_SMA_38 (t=5.97, p<.01, d=1.28) and between E2_SMA_38 and E2_SS_113 (t=7.35, p<0.01, d=1.34), but not between E2_SMC_42 and E2_SS_113 (t=0.23, p=.41, d=0.04).

approximately 78% - 91% of academic vocabulary types, while less than 3% of academic vocabulary types were shared by more than 75% of students and less than 8.5% of academic vocabulary types were shared by more than 50% of students. Thus, similar to the other sub-corpora, only a small number of academic types was shared by the majority of students. In other words, the diversity of academic vocabulary in Discussion Essays was, to a large extent, generated by a small proportion of students, suggesting that there was only a very small set of academic vocabulary types shared by the majority of students. The academic vocabulary types shared by the majority of the students are presented in Appendices 7f - 7h, containing 144 academic vocabulary items in total.

The Tables in Appendices 7f - 7h contain 45 (Appendix 7f), 70 (Appendix 7g) and 29 (Appendix 7h) academic vocabulary types, out of which 11 (24%), 17 (24%) and 6 (21%) vocabulary items (respectively), appeared in the assignment brief. Hence, the academic vocabulary types among those used by the majority of the students were formed to some extent (21% - 24%) by academic vocabulary contained in the assignment brief, pointing to the effect of the brief on the deployment of these academic vocabulary items, further discussed in 5.5.1.

Out of the 144 vocabulary items shown in Appendices 7f - 7h, 55 were found to be deployed in more than one sub-corpus of Discussion Essays (e.g. *provide*, *discuss*, *social*). These were counted once only, resulting in a core of 89 academic vocabulary items used in Discussion Essays. An analysis of the function of the 89 core academic vocabulary deployed by at least 50% of students showed that one third were research-oriented procedural items, approximately one quarter were research-oriented vocabulary items and around 22% of these vocabulary items were research-oriented vocabulary items relating to the assignment topic (Table 5.29).

Table 5.29: Academic vocabulary function (Discussion Essays)

Function	No (%)	Function	No (%)
RO-proc	30 (33.71%)	TO-trans	3 (3.37%)
RO-desc	22 (24.72%)	TO-res	3 (3.37%)
RO-top	20 (22.47%)	TO-fram	3 (3.37%)
TO-struc	7 (7.87%)	RO-loc	1 (1.12%)

The next section provides a discussion of the above results relating to the diversity of academic vocabulary across the four writing genres investigated in this study.

5.5 Diversity of academic vocabulary across writing genres: discussion

The analysis of the diversity of academic vocabulary in the four writing genres generated findings relating to the core academic vocabulary used by IFP students in their written assignments and the function of academic vocabulary deployed by the majority of the students. Similar to the density of academic vocabulary discussed above, it also highlighted the role of the topic, assignment brief and writing genre. These findings are further discussed below.

5.5.1 Core academic vocabulary

The diversity of academic vocabulary was explored from the perspective of the distribution of the identified academic vocabulary types among students to establish how the diversity of academic vocabulary was achieved. Table 5.30 below provides a summary of the number and percentage of academic vocabulary types used by ≥75%, ≥50%, ≥25%, <25% of students and individual students in the different sub-corpora as well as the average usage (M) of academic vocabulary types by the different proportions of students in each writing genre.

Table 5.30: Academic vocabulary types distribution

Exposition Essays					
Sub-corpus	**≥75% of SS**	**≥50% of SS**	**≥25% of SS**	**<25% of SS**	**Individual SS**
E1_LS_42	7 (0.9%)	30 (3.86%)	96 (12.36%)	681 (87.64%)	306 (39.38%)
E1_MI_38	5 (0.71%)	32 (4.56%)	108 (15.41%)	593 (84.59%)	258 (36.8%)
E1_FN_113	3 (0.34%)	20 (2.29%)	40 (4.58%)	834 (95.42%)	317 (36.27%)
Average	**M=0.65%**	**M=3.57%**	**M=10.78%**	**M=89.22%**	**M=37.48%**
Problem Questions					
Sub-corpus	**≥75% of SS**	**≥50% of SS**	**≥25% of SS**	**<25% of SS**	**Individual SS**
PQ_BOS_48	25 (2.43%)	73 (7.1%)	173 (16.83%)	855 (83.17%)	337 (32.78%)
PQ_ESP_73	26 (2.49%)	55 (5.27%)	138 (13.23%)	905 (86.77%)	347 (33.27%)
PQ_ST_59	48 (4.2%)	95 (8.32%)	220 (19.26%)	922 (80.74%)	353 (30.91%)
PQ_SI_13	30 (5.67%)	81 (15.31%)	187 (35.35%)	342 (64.65%)	260 (49.15%)
Average	**M=3.70%**	**M=9.00%**	**M=21.17%**	**M=78.83%**	**M=36.53%**
Research Reports					
Sub-corpus	**≥75% of SS**	**≥50% of SS**	**≥25% of SS**	**<25% of SS**	**Individual SS**
RR_193	**18 (1.14%)**	**66 (4.18%)**	**167 (10.58%)**	**1412 (89.42%)**	**394 (24.95%)**
Discussion Essays					
Sub-corpus	**≥75% of SS**	**≥50% of SS**	**≥25% of SS**	**<25% of SS**	**Individual SS**
E2_SMC_42	13 (1.48%)	45 (5.12%)	131 (14.9%)	748 (85.1%)	312 (35.49%)
E2_SMA_38	26 (3.06%)	70 (8.23%)	186 (21.86%)	665 (78.14%)	295 (34.67%)
E2_SS_113	7 (0.59%)	29 (2.45%)	104 (8.79%)	1079 (91.21%)	354 (29.92%)
Average	**M=1.71%**	**M=5.27%**	**M=15.18%**	**M=84.82%**	**M=33.36%**

Table 5.30 shows that overall the diversity of academic vocabulary was, to a large extent, achieved by a small number (<25%) of students: overall, approximately 79% - 89% of academic vocabulary types were generated by less than 25%, while less than 4% of academic vocabulary types were shared by more than 75% of students and around 4% - 9% were shared by half of the students. This suggests that in all writing genres there was only a very small set of core academic vocabulary types shared by the majority (i.e. at least 50%) of students in their written assignments. This finding is in line with Durrant's (2016) study investigating the extent to which university students draw on the AVL (3.2.5), which showed that "the majority of coverage was achieved by only a small number of items" (p. 59).

A closer examination of the distribution of the academic vocabulary types in individual writing genres showed differences in the distribution of the AVL items among the writing genres. Table 5.30 above shows that the Problem Question assignments generated the highest proportion of academic vocabulary types deployed by at least 75% of students (M=3.7%) compared to Exposition Essays (M=0.65%), Research Reports (1.14%) and Discussion Essays (M=1.75%). Similarly, the writing genre of Problem Questions showed the highest proportion of academic vocabulary types deployed by at least half of the students (M=9%) compared to Exposition Essays (M=3.57%), Research Reports (4.18%) and Discussion Essays (M=5.27%). This is also the case with academic vocabulary types used by at least 25% of students, where the Problem Questions generated the highest proportion of academic vocabulary types (M=21.17%) compared to Exposition Essays (M=10.78%), Research Reports (10.58%) and Discussion Essays (M=15.18%). This suggests that the writing genre of Problem Questions prompted a narrower repertoire of academic vocabulary, resulting in a higher number of academic vocabulary types being shared by a higher number of students than in the other writing genres. This finding hence points to the impact of the writing genre on the diversity of academic vocabulary in the students' assessed academic writing, further discussed in 5.5.3.

A further exploration of the core academic vocabulary showed the impact of the assignment briefs on the repertoire of academic vocabulary in the students' written production, where academic vocabulary items contained in the assignment brief formed 36% of all academic types in the Research Reports and up to 31% in the Exposition Essays. Hence, similar to the density of academic vocabulary, the role that

the assignment brief seemed to play in the diversity of academic vocabulary demonstrated in the student written assignments ought not to be overlooked as it appeared to be an important source of academic vocabulary for students, also found by others (e.g., Flowerdew, 2003; Milton, 2001) (3.2.2).

5.5.2 Core academic vocabulary function

The core academic vocabulary types deployed by the majority of students across the four writing genres were further explored in their co-textual environment from the perspective of their function in order to gain additional insights into the characteristics of these items (Table 5.31).

Table 5.31: Core academic vocabulary function

Ac. vocab. function	Exposition Essays	Problem Questions	Research Reports	Discussion Essays
RO-top	24 (41.38%)	28 (18.06%)	5 (7.58%)	20 (22.47%)
RO-proc	14 (24.14%)	52 (33.55%)	22 (33.33%)	30 (33.71%)
RO-desc	8 (13.79%)	48 (30.97%)	19 (28.79%)	22 (24.72%)
RO-loc	1 (1.72%)	1 (0.65%)	-	1 (1.12%)
TO-struc	6 (10.34%)	14 (9.03%)	15 (22.73%)	7 (7.87%)
TO-res	2 (3.45%)	4 (2.58%)	1 (1.52)	3 (3.37%)
TO-trans	2 (3.45%)	3 (1.94%)	2 (3.03%)	3 (3.37%)
TO-fram	1 (1.72%)	5 (3.23%)	2 (3.03%)	3 (3.37%)

Table 5.31 shows the impact of the topic in the Exposition Essays, where over 40% of the core academic vocabulary types related to the assignment topic. A possible explanation for this might be the social purpose of this writing genre to "develop the ability to construct a coherent argument and employ critical thinking skills" (Gardner & Nesi, 2013, p. 38) and the 'thesis - evidence - thesis restatement' structure characteristic of this writing genre requiring a series of arguments providing evidence to support the thesis and to argue for the stated position from the outset (Table 4.3 in 4.3.4), which may have prompted repeated references to the topic under discussion in all parts of the assignment (without an inclusion of alternative positions). This was evident from Figure 5.6 (in 5.2.1), for instance, which shows the usage of *harmful* (a topic-related academic vocabulary item drawn from the E1_FN_113 assignment title) throughout the assignment.

The topic-effect is least prevalent in the Research Reports. This is because even though the Research Reports related to the general topic of barriers to sustainability, the reports focused on a range of more specific topics selected by individual students (4.3.4). Problem Questions and Discussion Essays also displayed fewer research-

oriented topical academic vocabulary items. This may be explained by the fact that the Problem Question assignments aim to "provide practice in applying specific methods in response to professional problems" (Nesi & Gardner, 2012, p. 41), which required the usage of vocabulary relating to the procedures involved in these methods. This would also explain the relatively high proportion of research-oriented procedural academic vocabulary types in this writing genre. As for Discussion Essays, the lower proportion of topical academic vocabulary compared to Exposition Essays likely stemmed from the writing genre's requirement to include alternative positions, which may not have required the usage of academic vocabulary directly related to the topic under discussion.

As for research-oriented procedural academic vocabulary, despite approximately the same proportion of these items in Problem Questions, Research Reports and Discussion Essays, a closer inspection of these items showed that in Discussion Essays, the vast majority of the research-oriented procedural academic vocabulary were found to relate to general procedures (e.g. *enable*, *provide*, *act*, *promote*), while in Problem Questions and Research Reports some of these items related to the methodological aspects common in these two writing genres (e.g. *participate*, *research*, *conduct*, *observe*). This is because the Problem Questions were intended to give practice in the application of specific methods (noted above) and the Research Reports contained an entire section focusing on the description of methods (i.e. Introduction-Method-Results-Discussion, outlined in 4.3.4). Thus, the social purposes and the structure of the writing genres played a role in the characteristics (i.e. functions) of the academic vocabulary items deployed by the majority of students in Problem Questions and Research Reports.

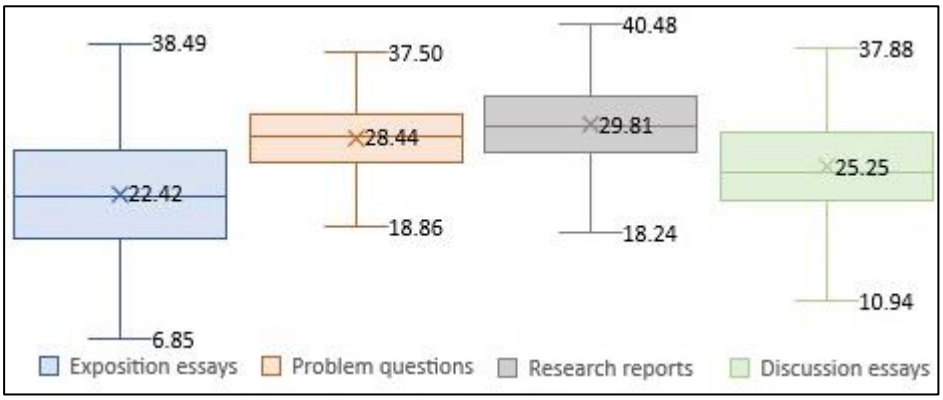
Further differences were found in the usage of text-oriented academic vocabulary. For example, text-oriented structuring academic vocabulary were most frequently deployed in Research Reports, probably due to the structure of this writing genre, which required the use of headings, figures and graphs (which were among the most frequent text-oriented structuring vocabulary items). Text-oriented resultative, transition and framing academic vocabulary as well as research-oriented location items all appeared with negligible frequencies between none and five occurrences representing less than 4% of the core academic vocabulary items across all writing genres and are thus not discussed further.

From the above-discussed differences in the functions of the academic vocabulary used by the majority of students, the assignment topic emerged as a factor impacting the type of academic vocabulary deployed by the majority of the students in Exposition Essays. This seemed to be prompted by the social purpose of this writing genre, confirmed by the other functional categories of the core academic vocabulary across the four writing genres, which also appeared to be impacted by the social purposes of the writing genres under study. The writing genre hence played an important role in the diversity of academic vocabulary and is further discussed next.

5.5.3 The role of the genre

The examination of the core academic vocabulary types used by the majority of students (Table 5.30 above) and their functional categories (Table 5.31 above) highlighted the effect of the writing genre on the differences in the way the core academic vocabulary types were used in IFP students' written assignments. This was shown by the impact of the social purposes of the writing genres on the function of the core academic vocabulary. Figure 5.27 (in 5.4), replicated in Figure 5.32 below, shows the differences in the diversity of academic vocabulary among the four writing genres investigated in this study (which were found to be statistically significant in section 5.4) with the Research Reports displaying the highest diversity of academic vocabulary.

Figure 5.32: Diversity of academic vocabulary (summary overview)

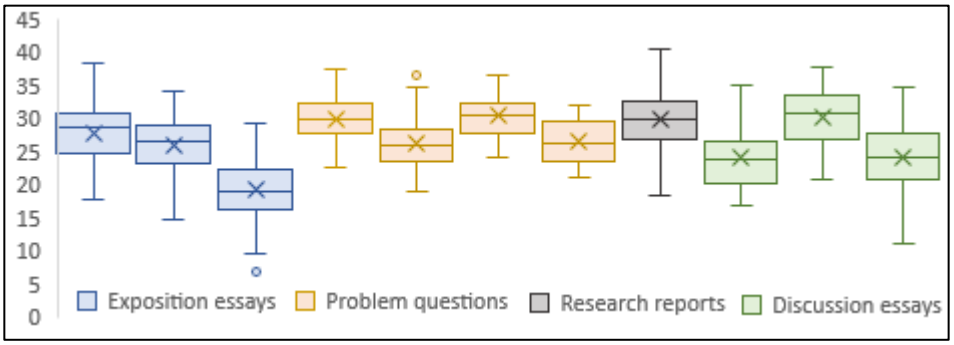


A possible explanation for the Research Reports showing the highest diversity of academic vocabulary is this writing genre's 'Introduction-Method-Results-Discussion' structure with each section serving a different purpose, which was reflected in the different and more diverse usage of academic vocabulary.

The social purposes of the writing genres were also found to impact the characteristics (i.e. function) of academic vocabulary types deployed across the writing genres under study (5.5.2). This finding is in line with other studies, which also found differences in academic vocabulary usage across writing genres (e.g., Csomay & Prades, 2018; Durrant, 2016; Olinghouse & Wilson, 2013) (3.2.5). Csomay and Prades (2018), for instance, found that some text types (i.e. writing genres) produced by students displayed a considerably higher percentage of academic words than others, which they argue provides "evidence that academic vocabulary is text type specific in student writing" (Csomay & Prades, 2018, p. 114). Their analysis showed that these differences are due to the rhetorical purposes of different genres which have a strong impact on the amount of academic vocabulary used in student writing.

An investigation of individual sub-corpora representing the different writing genres (Figures 5.28, 5.29 and 5.31) accompanied by one-way ANOVAs showed significant differences in the diversity of academic vocabulary within writing genres, summarised in Figure 5.33.

Figure 5.33: Diversity of academic vocabulary (summary overview)



The differences within writing genres (Figure 5.33) indicate that in addition to the writing genre, there were other factors impacting the diversity of academic vocabulary in the students' assessed academic writing, which may have had a greater impact on the diversity of academic vocabulary than the writing genre. The less significant effect of the writing genre (compared to other factors) on the deployment of academic vocabulary was also noted by Durrant (2016), whose findings showed that a genre is not as large a factor impacting the usage of AVL items as discipline (3.2.5). However, due to the generic nature of the IFP under study (which did not focus on specific academic disciplines), it is difficult to corroborate Durrant's (2016) findings. Nevertheless, the writing genre of Problem Questions focused on various

topics reflecting different subject areas and disciplines (e.g. business or psychology), which can be regarded as representative of different disciplines. Considering that significant differences were found among the four titles of the Problem Question assignments, it can be said that in the case of Problem Questions, the discipline (represented by the various topics in this study) played a more significant role in the diversity of academic vocabulary than the writing genre. This finding would thus align with Durrant's (2016) finding that discipline is a greater driver of variation than genre, but would contradict Csomay and Prades's (2018) findings that "text type and rhetorical purpose play a greater role in the percentage of academic words than merely topic selection", meaning that "the genre that students are asked to produce influences academic word use more than topic" (p. 114) (3.2.5).

The above-discussed findings suggest that when investigating academic vocabulary usage in a writing genre without exploring individual sub-corpora representing different topics separately, the genre seemed to play a significant role in the diversity of academic vocabulary. However, an analysis of individual sub-corpora within a writing genre (reflecting different topics, some of which relate to specific subject areas or disciplines) pointed to considerable differences within a writing genre. This indicates that factors other than the writing genre played a more important role in the diversity of academic vocabulary in IFP students' written assignments.

5.5.4 Diversity of academic vocabulary across writing genres: summary

The investigation of the usage of academic vocabulary from the perspective of diversity of academic vocabulary across the four writing genres of assessed academic writing showed that there was a small core of productive academic vocabulary types shared by the majority of students with the Problem Question writing genre prompting a narrower repertoire of academic vocabulary than the other writing genres. An exploration of the academic vocabulary types shared by the majority of students pointed to the impact of the assignment brief and highlighted the different functions that these academic vocabulary types display across genres, potentially resulting from the social purposes of individual writing genres. However, the genre effect was found to be less prominent than the impact that the assignment topic and discipline had on the diversity of academic vocabulary.

5.6 Academic vocabulary across writing genres: summary

This chapter has presented results and discussion in response to RQ1 (i.e. What are the density and diversity of academic vocabulary in international foundation students' assessed academic writing across writing genres?). The analysis of the deployment of academic vocabulary items across the four writing genres of IFP student assessed academic writing from the perspective of density and diversity of academic vocabulary showed that the usage of academic vocabulary varied across the genres with only a small core of academic vocabulary types deployed by the majority of students across all writing genres. In addition, the usage of academic vocabulary was found to be largely influenced by the genre's social purposes, which manifested itself in the various functions and meanings that these academic vocabulary items displayed in their co-textual environments in the different genres, and in the case of the Problem Question assignments across different topic areas representative of various disciplines. However, the genre-effect was found to be less significant than the assignment brief and topic.

The assignment brief was shown to play an important role in both the density as well as diversity of academic vocabulary as it was found to serve the function of academic vocabulary repository, particularly for less proficient novice student writers in English, who were found to draw on the vocabulary contained in the brief to a greater extent compared to those who did not report encountering difficulties with appropriate vocabulary and paraphrasing techniques. The assignment topic was found to be a further factor impacting the students' written production in terms of academic vocabulary deployment, highlighting the importance of not only knowledge of topic-specific vocabulary items, but also their meanings and collocational preferences in the specific contexts in which they are used due to the different behaviour that certain words display in different disciplines.

In summary, the investigation of academic vocabulary in response to RQ1 showed that the deployment of academic vocabulary items across the four writing genres of IFP students' assessed academic writing varied and was influenced by the assignment topic, assignment brief and the genre's social purposes.

Chapter 6: Development of academic vocabulary

6.1 Introduction

This chapter presents results relating to the development of academic vocabulary in the students' written assignments over the duration of the IFP (RQ2). The analysis of the development of academic vocabulary is approached from the perspective of changes in the density and diversity of academic vocabulary (4.4.3). For the purpose of measuring the changes in academic vocabulary usage, the genre family of Essays comprising Exposition (E1) and Discussion (E2) Essays was used, representing the first (E1) and last (E2) written assignment on the IFP collected for the purpose of this study (4.3.4). Table 6.1 provides a summary of the sub-corpora representing the two genres together with the assignment topics, number of texts forming each sub-corpus as well as the students' location and information relating to the size of the sub-corpora in terms of running words (i.e. tokens), average number of tokens per text, the required word length as stated in the assignment brief, the raw number of academic tokens and the percentage of academic tokens per all tokens, the raw number of types as well as the raw number of academic types and percentage of academic types per all types.

The changes in the density (6.2) and diversity (6.3) of academic vocabulary are investigated separately as follows: the results relating to the development of the density of academic vocabulary are focused on first (6.2.1), including an exploration of the potential factors impacting these changes (6.2.2 – 6.2.3). This is followed by the results relating to the development of the diversity of academic vocabulary (6.3.1), together with an investigation of the potential contributing factors affecting these changes (6.3.2 – 6.3.5). Since the analysis of the development of academic vocabulary from the perspective of density and diversity led to several common findings, the next section provides a discussion of the results relating to both density and diversity of academic vocabulary (6.4). Throughout these sections, visualisations are used to aid the presentation of results and their subsequent discussion. These are primarily in the form of boxplots, tables and figures containing extracts from the students' essays. This chapter concludes with a summary of the main findings relating to RQ2 (6.5).

Table 6.1: Assignments overview (RQ2)

Exposition Essays							
Sub-corpus (Assignment topic)	**No. of texts (campus)**	**Size (tokens)**	**Average tokens per text**	**Required word length[19]**	**No. of ac. tokens (% per tokens)**	**No. of types**	**No. of ac. types (% per types)**
E1_LS_42 (Learning styles)	42 (UK, MRU)	33,789	800	720 (+/- 10%)	5,649 (16.72%)	2,569	777 (30.25%)
E1_MI_38 (Multiple intelligences)	38 (UK, MRU)	35,326	930	720 (+/- 10%)	4,962 (14.05%)	2,656	701 (26.39%)
E1_FN_113 (Fake news)	113 (UAE)	77,163	680	650 – 750	8,972 (11.63%)	4,437	874 (19.7%)
Discussion Essays							
E2_SMC_42 (Social media and crime)	42 (UK, MRU)	55,815	1,330	1,350 (+/- 10%)	9,054 (16.22%)	4,269	879 (20.59%)
E2_SMA_38 (Social media for academic purposes)	38 (UK, MRU)	51,355	1,350	1,350 (+/- 10%)	10,896 (21.22%)	3,338	851 (25.49%)
E2_SS_113 (Surveillance society)	113 (UAE)	132,991	1,180	1,200 - 1,500	19,252 (14.48%)	5,746	1,183 (20.59%)

[19] as stated in the assignment brief in the Assessment Handbook

6.2 Density of academic vocabulary: results

6.2.1 Changes in the density of academic vocabulary

This section focuses on the density of academic vocabulary in Exposition (E1) and Discussion (E2) essays (Table 6.1) and explores the changes between them from the perspective of an increase or decrease in the density of academic vocabulary (Figure 6.1).

Figure 6.1: Exposition and Discussion Essays (density of academic vocabulary)

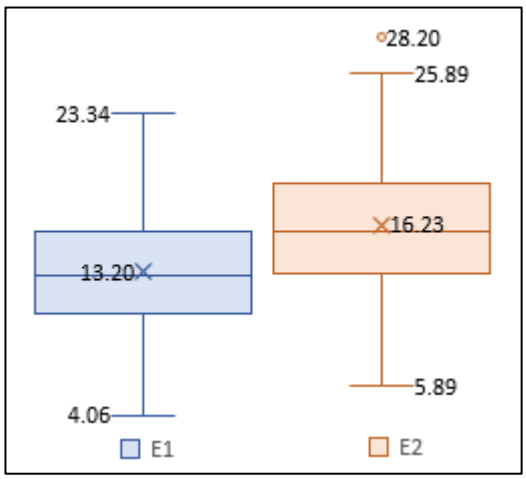


Figure 6.1 shows an overall increase in the density of academic vocabulary from E1 (with a mean average of 13.2%) to E2 (with a mean average of 16.23%), which was found to be statistically significant [20]. A closer examination revealed that 149 students (77%) showed a higher density of academic vocabulary in E2. This increase ranged between 0.13 - 16.49 percentage points (M=4.63 percentage points). The 44 students (23%) whose density of academic vocabulary did not increase showed a decrease of 0.2 - 8.73 percentage points (M=2.4 percentage points).

The highest increase in the density of academic vocabulary was demonstrated by Student UK15/16/G, whose density of academic vocabulary increased from 7.96% in E1 to 24.45% in E2 (an increase of 16.49 percentage points). Figures 6.2 and 6.3 below show an essay extract from this Student's E1 and E2.

[20] A paired sample t-test confirmed that the difference between E1 and E2 was statistically significant with a large effect size: t(192)=9.98, p<.00001, d=0.78.

Figure 6.2: Student UK15/16/G density of academic vocabulary E1 (E1_MI_38)

Howard Gardner was a Professor of Cognition and Education at the Harvard Graduate School of Education and the protagonist of the multiple intelligence theory. He believed that intelligence could not be only measured by mathematical or linguistic tests. Therefore he defines intelligence to be not just a simple grade from the spoon-fed education in other worlds, the pen and paper does not define one's intelligence. Apart the intelligence worldwide know, mathematics, logistic, and linguistics, which can be measured, Gardner also point out other kind of intelligences that he says to be "biopsychological". To put it more simply, those are the kind of intelligence that can't be measured by those tests. For example Wolfgang Amadeus Mozart was considered a genius in music but did not finish his studies in Law and Philosophy. With this in mind he was not supposed to be a genius, because the intelligence test will not show his ability towards music or his other abilities. Similarly this strange fact also applies to athletes, chess players, painter, etc. To sum up briefly, Gardner's theory indicates that intelligence is not only the ones we can size with a test but also the ones outlined above.

Figure 6.2 shows a somewhat scarce usage of academic vocabulary items (compared to Figure 6.3 below). In addition, repetition of certain academic lemmas (e.g. *measure x3* or *theory x2*) is evident. A number of the academic vocabulary items can also be traced back to the assignment brief and / or title (e.g. *multiple*, *theory*, *example*). A further exploration of this Student's E1 showed that 22% (12/54) of all academic tokens was formed by academic vocabulary drawn from the assignment brief.

Figure 6.3: Student UK15/16/G density of academic vocabulary E2 (E2_SMA_38)

Dalsgaard found out how can relate social networking by academic purposes, and how social networking can be useful within university education by students. It was stated by Mason, that social networking gives students the possibility to share information and resources that are created by them and also provide the opportunity to others to have access by putting on platforms such as bookmarks, references, links, and notes. Nevertheless, it was argued by Mason that social networking sites cannot be considered the new Learning Management Systems. However, it was mentioned by the researcher that the social networking sites can be used by students for academic purposes such as discussion forums in order to improve the direct communication and collaboration that will reflect on students experience at university. Moreover, the SNS will provide students with a secondary support between them by sharing of information, ideas, thoughts, writings and notes. It can be concluded that SNS will offer a positive academic experience between students around the world and also develop their consciousness and awareness of the activities of other students.

Figure 6.3 illustrates denser usage of academic vocabulary compared to E1 (Figure 6.2). Some academic vocabulary items were repeated (e.g. *social networking* x5) and some could also be traced back to the assignment brief and / or topic (e.g. *social*, *university*, *experience*). Overall, 34% (101/301) of all academic vocabulary tokens

were formed by academic vocabulary contained in the assignment brief (an increase of 11 percentage points compared to E1).

The second highest academic vocabulary density increase was demonstrated by Student MRU18/F, whose density of academic vocabulary increased from 9.93% in E1 to 24.6% in E2 (an increase of 14.67 percentage points). An examination of the Student's E1 and E2 showed that the increase was also impacted by academic vocabulary items drawn from the assignment brief; these academic vocabulary items formed 15% (23/149) of all academic tokens in E1 and 32% (102/320) in E2 (an increase of 17 percentage points). This points to the impact of the assignment brief on the increased density of academic vocabulary in E2. The analysis of the two students' essays with the highest increase in the density of academic vocabulary thus shows that this increase was partly attributed to drawing on academic vocabulary contained in the assignment brief to a greater extent in E2 than in E1.

The biggest decrease in the density of academic vocabulary was from 23.34% in E1 to 14.61% in E2 (a decrease of 8.73 percentage points) (Student MRU17/18/U). The less dense usage of academic vocabulary in E2 compared to E1 is illustrated by the essay extracts in Figures 6.4 and 6.5.

Figure 6.4: Student MRU17/18/U density of academic vocabulary E1 (E1_LS_42)

The accommodator is described by Kolb as someone who gathers information in order to resolve problems. In order to decipher a solution, accommodators analyse all the information they have gathered to adopt the best approach. For instance, I studied Business Studies for my A-level, which greatly contributed in improving my problemsolving skills. Furthermore, by learning the family business, reading previous case studies and through research, I was able to quickly and analytically provide a strategic answer. By fully involving myself in the situation and relating it to my current life, I am able to intuitively find a solution.

Figure 6.5: Student MRU17/18/U density of academic vocabulary E2 (E2_SMC_42)

Bullying was likely to use instant messaging, blogs and chat rooms on venues such as Facebook, Twitter, Instagram amongst others. The case of Amanda Todd is a good illustration of epidemic; she was a teenager who through Facebook had a nude picture of her being circulated rapidly. She was brutally cyberbullied leading her to commit suicide regardless, she was still bullied for her suicide. Illustrating the cruelty of some individuals and how the anonymity of the internet has expedited the act. Correspondingly, victims suffered depressive symptoms, excessive psychosomatic behavior, and neuroticism as written by Ybarra and Mitchell. The majority of victims are too ashamed to seek help, thus aggrading the issue.

A closer examination of the essays showed that even though Student MRU17/18/U demonstrated the highest decrease in the density of academic vocabulary, there was a slight increase in the deployment of academic vocabulary contained in the assignment brief: in E1, academic vocabulary items from the assignment brief formed 14% (26/186) of all academic items compared to E2 where academic vocabulary items drawn from the assignment brief formed 17% (28/169) of all academic vocabulary items. Hence, the overall decrease in the density of academic vocabulary in the Student's E2 was not impacted by a lower density of academic vocabulary items drawn from the assignment brief.

The second highest decrease in the density of academic vocabulary was found in Student UAE17/18/XXX's essay, where the density decreased from 15.98% in E1 to 8.23% in E2 (a decrease of 7.75 percentage points). Similar to Student MRU17/18/U, although there was a decrease in the density of academic vocabulary, an examination of the Student's essays showed that the usage of academic vocabulary items contained in the assignment brief per all academic items increased from 5% (5/93) in E1 to 12% (11/89) in E2.

The texts of the students who showed both an increase as well as a decrease in the density of academic vocabulary closest to the average (increase M=4.63 percentage points; decrease M=2.4 percentage points) were also explored. An examination of the students' essays with an average increase in the density of academic vocabulary[21] showed that only one student (UAE17/18/LL) demonstrated an increase in the usage of academic vocabulary items contained in the assignment brief, which increased from 6% (6/98) in E1 to 16% (34/209) in E2. The other two students' usage of academic vocabulary drawn from the assignment brief decreased by 2 percentage points (from 14% in E1 to 12% in E2) and 5 percentage points (from 43% in E1 to 38% in E2).

Similarly, an examination of the essays of the students demonstrating a decrease in the density closest to the average[22] showed both an increase and decrease in the

[21] These students were: Student UAE/17/18/OOO who demonstrated an increase of 4.71% (from 4.06% in E1 to 8.77% in E2); Student UK16/17/D whose density increased by 4.64% (from 11.5% in E1 to 16.14% in E2); and Student UAE17/18/LL with an increase of 4.57% (from 13.35% in E1 to 17.92% in E2).

[22] Student UK14/15/F's density of academic vocabulary decreased by 2.29% (from 13.32% in E1 to 11.03% in E2) and Student MRU17/18/G showed a decrease of 2.61% (from 18.64% in E1 to 16.03% in E2) in the density of academic vocabulary.

deployment of academic vocabulary items contained in the assignment brief where one of the Students (UK14/15/F) displayed a decrease of the usage of academic vocabulary items drawn from the assignment brief by 5 percentage points (from 30% in E1 to 25% in E2), whereas the other Student (MRU17/18/G) showed an increase in the usage of academic vocabulary items contained in the assignment brief by 23 percentage points (from 5% in E1 to 28% in E2).

In sum, the investigation into the changes in the density of academic vocabulary exemplified by the students with the highest increase and decrease showed that in both instances there was a denser usage of academic vocabulary items drawn from the assignment brief in E2 compared to E1. However, an examination of a number of students' texts with an average increase and decrase in the density of academic vocabulary showed mixed results. Thus, the usage of the academic vocabulary items contained in the assignment brief between E1 and E2 was investigated further.

6.2.2 The impact of the assignment brief

The impact of the assignment brief was explored next with the aim of examining a general trend in the IFP studetns' assessed writing. Table 6.2 shows the percentage of the academic vocabulary items drawn from the assignment brief per all academic tokens in each sub-corpus and indicates whether there was an increase (↑) or decrease (↓) from E1 to E2.

Table 6.2: Impact of the assignment brief on density (E1 and E2)

E1	E2	↑↓
E1_LS_42: 17.63% of ac. tokens	**E2_SMC_42:** 24.13% of ac. tokens	↑
E1_MI_38: 16.61% of ac. tokens	**E2_SMA_38:** 29.28% of ac. tokens	↑
E1_FN_113: 8.83% of ac. tokens	**E2_SS_113:** 15.16% of ac. tokens **M =**	↑
M = 14.36%	**22.86%**	↑

Table 6.2 shows that the proportion of academic tokens formed by academic vocabulary items drawn from the assignment brief increased overall (from 14.36% in E1 to 22.86% in E2) as well as in each sub-corpus. This may suggest that the students developed the strategy of drawing on vocabulary items contained in the assignment brief as a way of addressing the assignment instructions. The assignment brief hence seemed to be a contributing factor impacting the changes in the deployment of academic vocabulary, whereby it led to increased density of academic vocabulary in the students' written assignments. This finding is also corroborated by the survey

data (Chapter 7; 7.3.1), according to which nearly 40% of students acknowledged that the assignment brief had played an important role in the deployment of academic vocabulary in the process of completing their written assignments. It is also noteworthy that all E2 assignment briefs contained a higher number of academic vocabulary items than titles under E1 (Table 6.3), which may have impacted the denser usage of these academic vocabulary items in E2, further pointing to the important role that the vocabulary items contained in the assignment brief played in the students' written production.

Table 6.3: Academic vocabulary items contained in assignment briefs (E1 and E2)

Writing genre	Sub-corpus	No. of tokens	No. of ac. lemmas in brief	% of ac. lemmas from brief per ac. tokens	% of ac. lemmas from brief per all tokens
Exposition Essays (E1)	E1_LS_42	165	26	17.63%	2.95%
	E1_MI_38	156	22	16.61%	2.33%
	E1_FN_113	110	18	8.83%	1.03%
Discussion Essays (E2)	E2_SMC_42	229	41	24.13%	3.91%
	E2_SMA_38	248	49	29.28%	6.21%
	E2_SS_113	215	32	15.16%	2.12%

This finding thus suggests that over the duration of the IFP the students developed the strategy of integrating the academic vocabulary contained in the assignment brief in their written assignments with an increased frequency, resulting in an increased density of these academic vocabulary items in their writing.

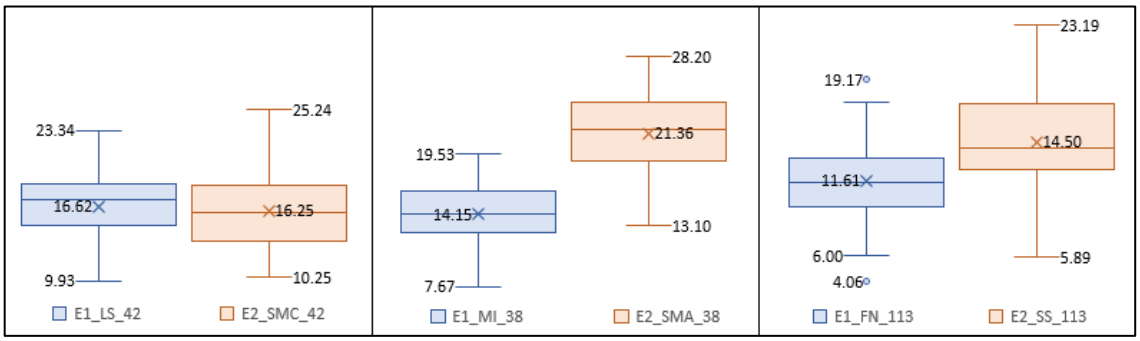
The importance of the assignment brief was also noted in response to RQ1 (Chapter 5), where it was hypothesised that the reliance on the vocabulary items contained in the assignment brief may have been a compensatory strategy of novice writers who were less confident in using their own words or those who found the usage of appropriate vocabulary difficult, who may have thus used the brief as a language repository (5.3.1). Nevertheless, as noted above, the denser usage of academic vocabulary items contained in the assignment brief in E2 compared to E1 may also suggest that drawing on the assignment brief was not only a compensatory strategy for less proficient and / or less confident novice writers in English, but that the students developed the strategy of addressing the assignment instructions more explicitly by integrating these academic vocabulary items in their assignments.

6.2.3 The impact of the topic

A further investigation into the changes in the density of academic vocabulary between E1 and E2 showed that 62% of the students (92 out of 149) who demonstrated an increase in the density of academic vocabulary were those who completed E1 on Fake News (E1_FN_113) and E2 on Surveillance Society (E2_SS_113), followed by 25% of students (37) whose E1 was on Multiple Intelligences (E1_MI_38) and E2 on Social Media for Academic Purposes (E2_SMA_38). Only 13% of students (20) whose E1 related to Learning Styles (E1_LS_42) and E2 to Social Media and Crime (E2_SMC_42) demonstrated an increase in the density of academic vocabulary.

The higher number of students from the Fake News (E1_FN_113) and Surveillance Society (E2_SS_113) sub-corpora is perhaps unsurprising considering that these two sub-corpora were formed by the biggest group of students (113) compared to the other two sub-corpora (formed by 42 and 38 students). However, a further examination showed that among the students who demonstrated an increase in the density of academic vocabulary of at least 10 percentage points were students predominantly from the E1_MI_38 / E2_SMA_38 sub-corpora (9 out of 11 students) and that texts from the E1_LS_42 / E2_SMC_42 sub-corpora all demonstrated an increase in the density of academic vocabulary no higher than 7 percentage points. This prompted the question of the impact of the assignment topic on the changes in the density of academic vocabulary, further investigated next by comparing pairs of sub-corpora completed by the same students (Figure 6.6). These pairs of sub-corpora were formed by 42 students (E1_LS_42 & E2_SMC_42), 38 students (E1_MI_38 & E2_SMA_38) and 113 students (E1_FN_113 & E2_SS_113).

Figure 6.6: Exposition and Discussion Essays (density of ac. vocab. in paired sub-corpora)



As can be seen from Figure 6.6, the density of academic vocabulary decreased in the first pair of sub-corpora, while there was an increase in the density of academic vocabulary in the other two pairs of sub-corpora, which was found to be statistically significant[23]. This supports the above finding relating to the highest increase in the density of academic vocabulary in the E1_MI_38 / E2_SMA_38 pair of sub-corpora. This also matches the findings relating to RQ1 (Chapter 5) which showed the impact of the assignment topic on the density of academic vocabulary (5.3.2).

This is particularly interesting in the case of E2_SMC_42 (Social Media and Crime) and E2_SMA_38 (Social Media for Academic Purposes) pair of sub-corpora given the topical similarities between them (i.e. the focus on social media) and the large difference in the changes in density of academic vocabulary between these two sub-corpora. This indicates that even subtle topical variations may have impacted the density of academic vocabulary. Hence, this finding points to the association between the influence of the topic and the potential changes in the density of academic vocabulary, whereby some topics seemed to prompt denser usage of academic vocabulary items than others.

6.2.4 Density of academic vocabulary: summary

In sum, the above analysis of the changes in the density of academic vocabulary over the duration of the IFP has shown that there was an overall increase from E1 to E2 with more than three quarters of students (77%) demonstrating an increase in the density of academic vocabulary with a mean average of 4.63 percentage points. A closer investigation showed that in all sub-corpora the increase was achieved, to some extent, by increased density of academic vocabulary items drawn from the assignment brief. A further examination of individual sub-corpora showed that only two out of the three pairs of sub-corpora showed an increase in the density of academic vocabulary. This indicated that some topics may have prompted a denser usage of academic vocabulary than others.

[23] A paired sample t-test generated the following results for the three pairs of sub-corpora displayed in Figure 6.6: E1_LS_42 & E2_SMC_42: t(41)=-0.72, p=.476, d=0.12; E1_MI_38 & E2_SMA_38: t(37)=11.73, p<.00001, d=2.25; E1_FN_113 & E2_SS_113: t(112)=9.02, p<.00001, d=0.92.

6.3 Diversity of academic vocabulary: results

6.3.1 Changes in the diversity of academic vocabulary

This section presents findings relating to the diversity of academic vocabulary in Exposition (E1) and Discussion (E2) Essays and investigates the changes between these two essays with regard to the increase or decrease in the diversity of academic vocabulary (Figure 6.7).

Figure 6.7: Exposition and Discussion Essays (diversity of academic vocabulary)

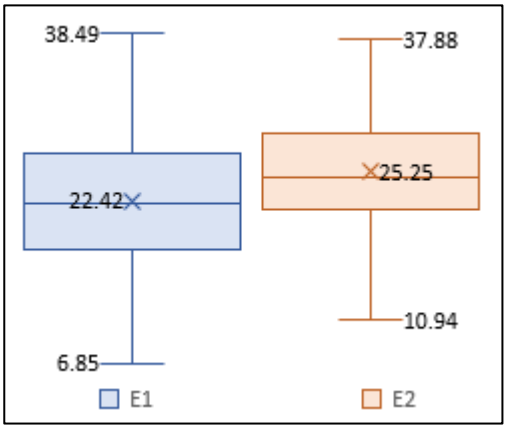


Figure 6.7 shows an overall increase in the diversity of academic vocabulary from E1 (with a mean average of 22.42%) to E2 (with a mean average of 25.25%), which was found to be statistically significant[24]. A further analysis showed that 70% of students (135 out of 193) demonstrated an increase in the diversity of academic vocabulary in E2. This increase ranged between 0.14 - 18.99 percentage points (M=5.89 percentage points). The 58 students (30%) whose diversity of academic vocabulary decreased showed a decline between 0.2 - 14.91 percentage points (M=4.3 percentage points).

The highest increase in the diversity of academic vocabulary was demonstrated by Student UK15/16/G, whose diversity of academic vocabulary increased from 14.72% in E1 to 33.71 % in E2 (an increase of 18.99 percentage points). This Student also demonstrated the highest academic vocabulary density increase (Figures 6.2 and 6.3 in 6.2.1). Since the assignment brief was found to play an important role in the increase of the density of academic vocabulary not only in Student UK15/16/G's E2 but also overall (Table 6.2 in 6.2.2), its impact on the increase in the diversity of academic vocabulary was also investigated.

[24] A paired sample t-test confirmed that the difference between E1 and E2 was statistically significant with a medium effect size: t(192)=6.57, p<.00001, d=0.51.

6.3.2 The impact of the assignment brief

The analysis of Student UK15/16/G's essay, which showed the highest increase in the diversity of academic vocabulary, showed that this increase was only marginally (by approximately 2 percentage points) impacted by academic vocabulary items drawn from the assignment brief; that is, 18% (7/39) of academic vocabulary types were drawn from the assignment brief in E1, compared to 20% (24/118) in E2. The second highest increase in the diversity of academic vocabulary was demonstrated by Student UAE17/18/DDD, whose diversity increased from 13.85% in E1 to 29.98% in E2 (an increase of 16.13 percentage points). A further exploration showed that this increase was not impacted by the academic vocabulary contained in the assignment brief. In this Student's E1, these academic vocabulary items formed 10% (4/41) of all academic types compared to 6% (9/149) in E2.

The biggest decrease in the diversity of academic vocabulary was found in Student MRU17/18/U's E2, which showed a decline of 14.91 percentage points (from 38.49% in E1 to 23.58% in E2). This Student was also deploying fewer academic vocabulary items from the assignment brief in E2 (6/108; 6% of all academic types) compared to E1 (10/122; 8% of all academic types). Student MRU17/18/T, who demonstrated the second biggest decrease in the diversity of academic vocabulary (from 33.02% in E1 to 22.87% in E2; a decrease of 10.15 percentage points) drew on the academic vocabulary items from the assignment brief to a similar extent in E1 and E2 (11/94; 12%) and E1 (13/107; 12%).

The texts of the students who showed both an increase as well as a decrease in the diversity of academic vocabulary closest to the average (i.e. increase M=5.89 percentage points; decrease M=4.3 percentage points) were also explored. An examination of the students with an average increase in the density of academic vocabulary[25] showed that two students (UAE17/18/BB and UAE17/18/H) demonstrated an increase in the usage of academic vocabulary types contained in the assignment brief by 4 percentage points (from 7% in E1 to 11% in E2) and 1 percentage point (from 6% in E1 to 7% in E2) respectively, while one student

[25] These students were: Student UAE18/T who demonstrated an increase of 6.01% (from 16.74% in E1 to 22.75% in E2); Student UAE17/18/BB whose density increased by 5.9% (from 16.79% in E1 to 22.69% in E2); and Student UAE17/18/H with an increase of 5.86% (from 22.02% in E1 to 27.88% in E2).

(UAE18/T) showed a decrease of academic vocabulary types drawn from the assignment brief in E2 by 5 percentage points (from 16% in E1 to 11% in E2)

The students who showed a decrease in the diversity of academic vocabulary closest to the average[26] demonstrated a decrease of 1 percentage point (from 10% in E1 to 9% in E2, Student MRU17/18/R) and an increase of 2 percentage points (from 9% in E1 to 11% in E2, Student MRU17/18/I) in the usage of academic vocabulary types drawn from the assignment brief.

The analysis of the students showing the highest increase and decrease as well as an average increase and decrease in the diversity of academic vocabulary shows that although some students' increase in the diversity of academic vocabulary could be attributed to academic vocabulary types drawn from the assignment brief, this increase was only marginal. Hence, the assignment brief did not seem to impact the diversity of academic vocabulary to the same extent as the density of academic vocabulary (6.2.2); that is, the overall increase in the diversity of academic vocabulary was not caused by drawing on academic vocabulary items contained in the assignment brief. Table 6.4 shows that overall there was a slight decrease in the proportion of the academic vocabulary types drawn from the assignment brief per all types in E2.

Table 6.4: Impact of the assignment brief on diversity (E1 and E2)

E1	E2	↑↓
E1_LS_42: 26.67% of ac. types	**E2_SMC_42:** 24.44% of ac. types	↓
E1_MI_38: 31.25% of ac. types	**E2_SMA_38:** 24.29% of ac. types	↓
E1_FN_113: 15% of ac. types	**E2_SS_113:** 20.69% of ac. types	↑
M=25.61%	**M=23.61%**	**↓**

6.3.3 The role of reading sources

The above result relating to the assignment brief not playing an integral role in the increase in the diversity of academic vocabulary may be explained by the fact that towards the end of the IFP the students were drawing on a greater number of reading sources which may have served as a vocabulary repository, resulting in a larger repertoire of academic vocabulary and inclusion of academic vocabulary items other than those contained in the assignment brief. This is also supported by an exploration

[26] These were Student MRU17/18/R whose diversity decreased from 26.14% in E1 to 22.03% in E2 (a decrease of 4.11%), and Student MRU17/18/I who showed a decrease of 4.47% from 30.69% in E1 to 26.22% in E2.

of the number of reading sources on which the students drew in E1 and E2, which showed that on average, students were drawing on 7 reading sources in E1 in contrast to 14 sources in E2. This increase was found to be statistically significant with a large effect size (as confirmed by a paired sample t-test: $t(192)=15.67$, $p<.00001$, $d=1.17$).

The impact of reading sources is further corroborated by the survey as well as interview data. Out of the 51 students who completed the survey, nearly half (23 students, 45%) selected reading as one of the main perceived contributors impacting the changes in their academic writing; almost two-thirds (32 students, 63%) stated that seeing examples in sources had contributed to their perceived changes in academic writing, and a little less than one-third (15 students, 29%) of students reported that they had used vocabulary from texts as a writing strategy during the process of completing their written assignments. The survey data thus support the link between the use of reading sources and potential changes in the deployment of academic vocabulary in the students' writing.

This finding was further corroborated by the interview data. Student MRU18/F, for instance, reported in the interview that they had noticed changes in their academic writing and also pointed out the use of sources, which may have been a contributing factor resulting in the increased usage of academic vocabulary:

> "I've changed the way I write during the process of the IFP … the use of books, journals, articles, websites to do more research on the topic…" (Student MRU18/F).

The Student also showed awareness of the impact of reading on their vocabulary development by stating that they would continue to work on their vocabulary development by reading more.

The importance of reading sources was also noted by other students:

> "I've learned new words [from] the studies we would be reading, the examples of writing, I'd see a word and think I don't know this word, I need to know this word.. I'm gonna go and look it up because I need to know this word… so the reading, the research" (Student UK16/17/B).

> "What I have learned was through the journal articles that I looked through while researching" (Student UAE17/18/A).

"The articles, like journal articles and whatever else I was reading, that also helped a lot in improving the vocabulary I was using" (Student UAE18/K).

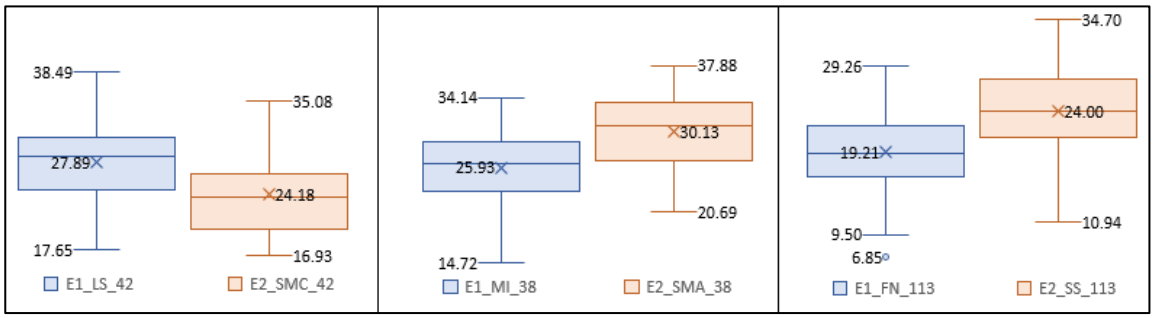
These findings thus highlight the important role that reading sources played in the deployment of academic vocabulary in the student writing production.

6.3.4 The role of the topic

A further analysis of the changes in the diversity of academic vocabulary between E1 and E2 showed that 71% of the students who demonstrated an increase (96 out of 135) were from the E1_FN_113 / E2_SS_113 pair of sub-corpora, followed by 22% of students (30) from the E1_MI_38 / E2_SMA_38 pair of sub-corpora. Only 7% of students (9) form the E1_LS_42 / E2_SMC_42 pair of sub-corpora showed an increase in the diversity of academic vocabulary. Out of the 58 students whose diversity of academic vocabulary decreased, more than half (33 students; 57%) were from this pair of sub-corpora. Further, no students from the E1_LS_42 / E2_SMC_42 pair of sub-corpora were among those showing an increase of at least 10%.

Similar to the density of academic vocabulary (6.2.3), the considerably lower number of students demonstrating an increase in the diversity of academic vocabulary from the E1_LS_42 / E2_SMC_42 sub-corpora prompted the question of the impact of the assignment topic on the changes in the diversity of academic vocabulary, investigated further by comparing paired sub-corpora of E1 and E2 (Figure 6.8).

Figure 6.8: Exposition and Discussion Essays (diversity of ac. vocab. in paired sub-corpora)



There was a statistically significant decrease in the first pair of sub-corpora, while a statistically significant increase was found in the two other pairs of sub-corpora[27].

[27]A paired sample t-test generated the following results for the three pairs of sub-corpora displayed in Figure 6.8: E1_LS_42 & E2_SMC_42: t(41)=-5.25, p<.00001, d=0.82; E1_MI_38 & E2_SMA_38: t(37)=4.54, p=.00006, d=0.96; E1_FN_113 & E2_SS_113: t(112)=10.83, p<.00001, d=1.05.

This corroborates the above findings relating to students from the E1_FN_113 / E2_SS_113 pair of sub-corpora demonstrating the highest increase in the diversity of academic vocabulary. Hence, although there was an overall increase in the diversity of academic vocabulary from E1 to E2, this increase was found in two of the three sub-corpora only.

As with the density of academic vocabulary (6.2.3), this finding may point to the impact of the assignment topic on the diversity of academic vocabulary as it suggests that some topics prompted a more diverse usage of academic vocabulary. This can be seen from the differences in the diversity of academic vocabulary in the E2 assignments on Social Media and Crime (E2_SMC_42) and Social Media for Academic Purposes (E2_SMA_38) (Figure 6.8), which illustrates that even subtle topical variations may have impacted the diversity of academic vocabulary. This finding relates to Laufer and Nation's (1995) study on lexical richness (i.e. diversity), in which they pointed out that one of the factors that may impact lexical richness in writing is familiarity with the topic, meaning that "a change of topic could result in a marked change in lexical richness" (p. 308). The findings of this study thus suggest that this may also have applied to the diversity of academic vocabulary in this study.

6.3.5 Core academic vocabulary

The changes in the size of the core academic vocabulary were investigated next by comparing the size of core academic vocabulary in individual pairs of sub-corpora of E1 and E2 as well as overall (Table 6.5).

Table 6.5: Core academic vocabulary (E1 and E2)

E1/E2	Sub-corpus	≥75% of SS	≥50% of SS	≥25% of SS	<25% of SS	Individual SS
E1	E1_LS_42	7 (0.9%)	30 (3.86%)	96 (12.36%)	681 (87.64)	306 (39.38)
	E1_MI_38	5 (0.71%)	32 (4.56%)	108 (15.41%)	593 (84.59%)	258 (36.8%)
	E1_FN_113	3 (0.34%)	20 (2.29%)	40 (4.58%)	834 (95.42%)	317 (36.27%)
	Average	**M=0.65%**	**M=3.57%**	**M=10.78%**	**M=89.22%**	**M=37.48%**
E2	E2_SMC_42	13 (1.48%) ↑	45 (5.12%) ↑	131 (14.9%) ↑	748 (85.1%) ↓	312 (35.49%) ↓
	E2_SMA_38	26 (3.06%) ↑	70 (8.23%) ↑	186 (21.86%) ↑	665 (78.14%) ↓	295 (34.67%) ↓
	E2_SS_113	7 (0.59%) ↑	29 (2.45%) ↑	104 (8.79%) ↑	1079 (91.21%) ↓	354 (29.92%) ↓
	Average	**M=1.71%** ↑	**M=5.27%** ↑	**M=15.18%** ↑	**M=84.82%** ↓	**M=33.36%** ↓

Table 6.5 shows that the size of the core academic vocabulary used by more than a quarter of students (i.e. ≥75%, ≥50%, ≥25%) increased from E1 to E2, while there was a decrease in academic vocabulary types generated by less than a quarter of students and individual students. It is interesting to note that both the increase as well as decrease in the size of core academic vocabulary was consistent across all pairs of sub-corpora for each proportion of students. This indicates that as the students progressed, they began to deploy a larger number of academic vocabulary types; in other words, there was a higher number of shared (i.e. core) academic vocabulary types among the students.

A further investigation into the academic vocabulary types shared by the majority of students (i.e. at least 50% of students) showed that this increase was achieved by academic vocabulary types displaying certain functions (Table 6.6).

Table 6.6: Core academic vocabulary function (E1 and E2)

Function	E1	E2	↑	Function	E1	E2	↓
RO-proc	24.14%	33.71%	↑	RO-top	41.38%	22.47%	↓
RO-desc	13.79%	24.72%	↑	TO-struc	10.34%	7.87%	↓
TO-fram	1.72%	3.37%	↑	RO-loc	1.72%	1.12%	↓
				TO-res	3.45%	3.37%	↓
				TO-trans	3.45%	3.37%	↓

There was an increase in research-oriented procedural and descriptive academic vocabulary as well as in text-oriented framing academic vocabulary types (Table 6.6). A closer analysis showed that among the research-oriented procedural academic vocabulary used in E1 were predominantly verbs (e.g. *apply*, *describe*, *develop*, *identify*, *improve*, *perform*, *reflect*, *relate*) (Figure 6.9), whereas a number of the research-oriented procedural academic vocabulary used in E2 were nouns (e.g. *activity*, *communication*, *interaction*, *performance, usage*) (Figure 6.10). This suggests that as the students progressed, they developed the ability to use nominal style of writing including nominalisation in their texts, which is regarded as characteristic of academic discourse (2.2.3).

Figure 6.9: Research-oriented procedural academic vocabulary E1

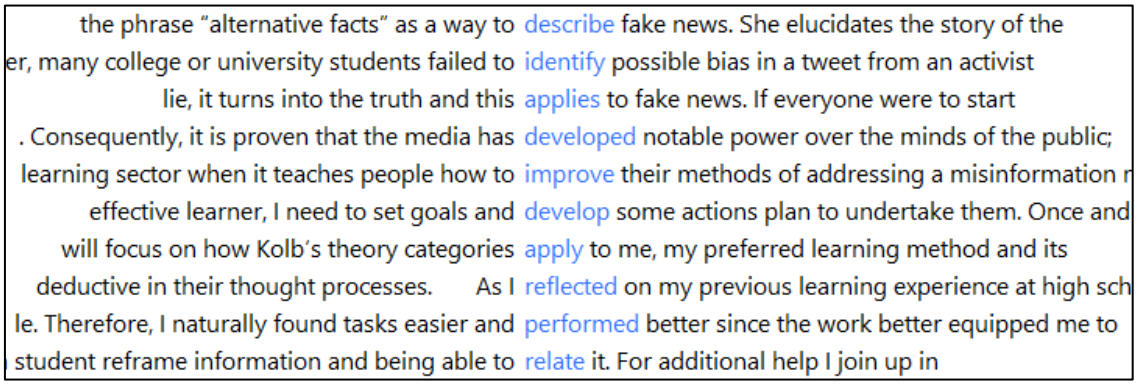
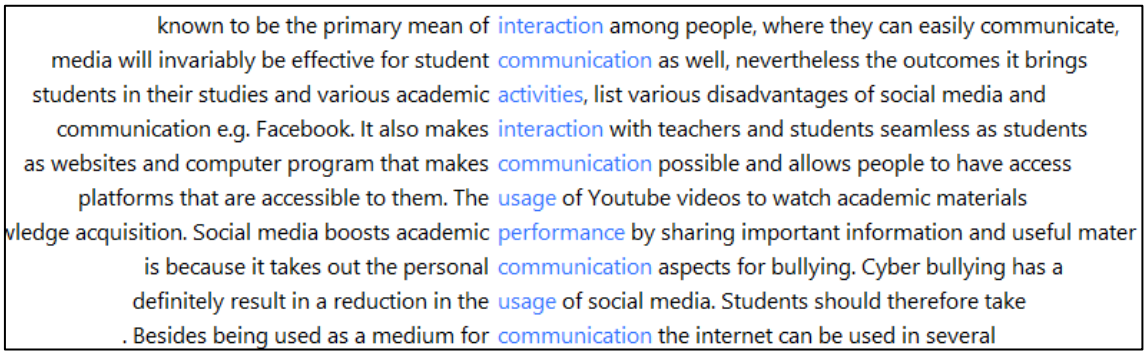


Figure 6.10: Research-oriented procedural academic vocabulary E2



As for the second-highest increase, found in research-oriented descriptive academic vocabulary, a further investigation showed an increase in research-oriented descriptive academic vocabulary represented by adjectives in E2 (e.g. *beneficial*, *important*, *positive*, *negative*), compared to E1 where the majority of these academic vocabulary items were formed by nouns (e.g. *characteristics*, *group*, *society*, *type*) (Figures 6.11 and 6.12).

Figure 6.11: Research-oriented descriptive academic vocabulary E1

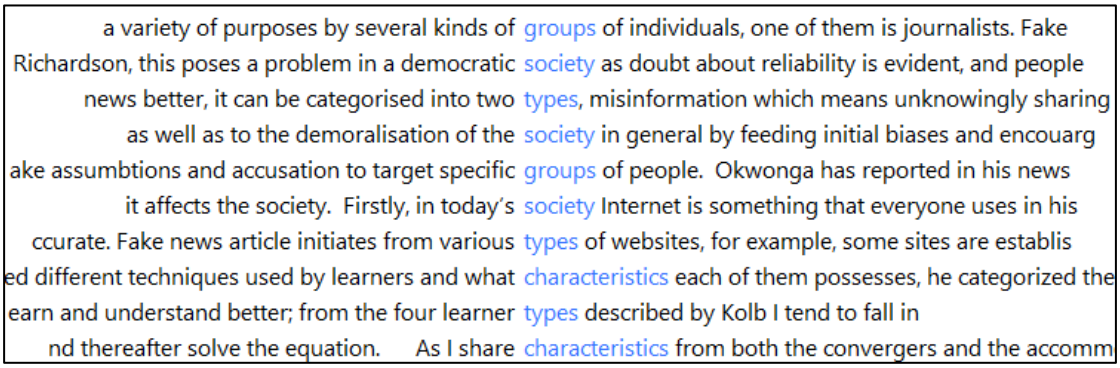
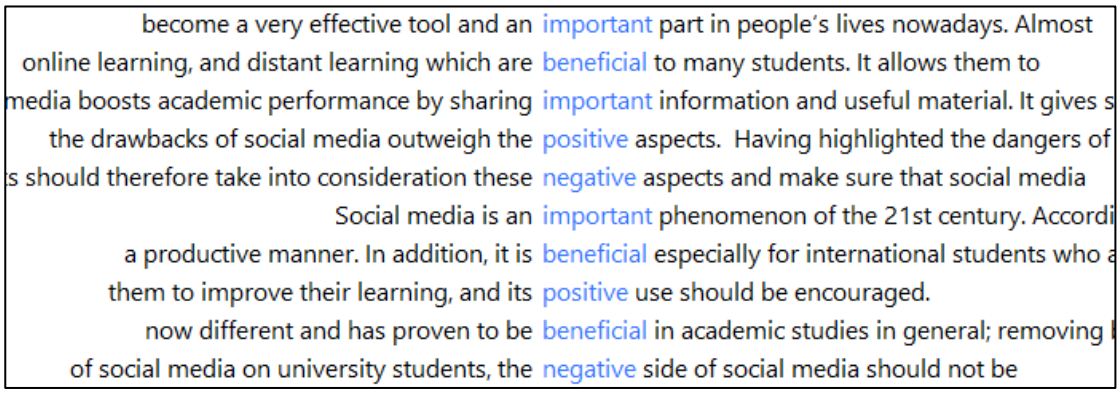


Figure 6.12: Research-oriented descriptive academic vocabulary E2



These examples of the descriptive academic vocabulary also suggest that the students began to incorporate evaluative expressions when presenting alternative positions and arguing for one of the positions, defined as "the expression of the speaker or writer's attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about" (Hunston & Thompson, 2000, p. 5). This indicates that the students started to incorporate academic vocabulary necessary to fulfil this genre's social purposes; that is, evaluation of evidence (Table 4.6 in 4.3.4). This can be regarded as an important development in the students' writing as "evaluative language is widely recognised as contributing to the quality of written argumentation" (Mei & Allison, 2005, p. 105) representing a crucial aspect of the writing genre the students were asked to produce (i.e. a Discussion Essay).

An increase, though only slight, was also identified in the usage of text-oriented framing signals. Among the text-oriented framing academic vocabulary deployed in E1 was *define* only, with *both* and *include* identified in addition to *define* in E2. *Both* was found to be used for the expression of different positions on the discussed issue (e.g. *both* advantages and disadvantages, *both* beneficial and detrimental, *both* for and against), whereby the genre's social purpose was fulfilled. *Include* was predominantly used to list examples relating to the issue under discussion, providing evidence for the points presented. These text-oriented framing academic vocabulary items can be categorised as metadiscourse, which is regarded as a defining feature of successful writing. Specifically, the identified framing academic vocabulary can be seen as relating to Hyland's (2005) metadiscourse categories of 'frame' markers (4.4.3).

The highest decrease was found in the usage of research-oriented topical academic vocabulary types (a decrease of approximately 18% from 41% in E1 to less than 23%

in E2). It is also interesting to note that the proportion of the text-oriented resultative and transition academic vocabulary remained approximately the same with a negligible decrease of 0.08 percentage points between E1 and E2. The approximately same level of usage of the resultative items may be explained by the fact that both Exposition and Discussion Essays required the writers to reach a position (4.3.4). In terms of the transition signals, a closer inspection of these academic vocabulary items revealed that only *furthermore* and *moreover* (expressing an addition) were identified among those academic vocabulary items used by the majority of students in E1. In E2, *however* (expressing a contrast) was added to these transition signals. This shows that the students were using academic vocabulary items to indicate an alternative position, as required by the E2 writing genre (i.e. a Discussion Essay) (Table 4.6 in 4.3.4), in contrast to E1 which required an inclusion of arguments supporting one position throughout (Table 4.3 in. 4.3.4).

These changes in the function of the core academic vocabulary (Tables 6.5 and 6.6) were likely to be impacted by the E2 writing genre's characteristics and could also be attributed to the students' developing certain writing strategies, which may have directly impacted the type of academic vocabulary items deployed in their written assignments (further discussed in 6.4.4).

6.3.6 Diversity of academic vocabulary: summary

In sum, the above analysis of the changes in the diversity of academic vocabulary showed an increase from E1 to E2 with 70% of students (135/193) demonstrating an increase in the diversity of academic vocabulary over the duration of the IFP with a mean average of 5.89 percentage points. A closer investigation showed that this increase was impacted by two out of the three pairs of sub-corpora. This suggests that certain topics may have prompted a more diverse usage of academic vocabulary than others. A further exploration showed that, unlike the increase in the density of academic vocabulary, the overall increase in the diversity of academic vocabulary was not caused by drawing on academic vocabulary items contained in the assignment brief. Instead, the more extended use of reading sources was likely to have contributed to the increased diversity of academic vocabulary. It was also found that the size of the core academic vocabulary types increased over the duration of the IFP. In addition, changes in the function of these core academic vocabulary types were identified. These findings are further discussed next.

6.4 Development of academic vocabulary: discussion

6.4.1 Changes in the density and diversity of academic vocabulary

The analysis of the development of the density and diversity of academic vocabulary in the IFP students' assessed academic writing over the duration of the IFP showed an increase in both the coverage or proportion that academic vocabulary items formed in the students' texts (i.e. density) (6.2) as well as in the repertoire of academic vocabulary (i.e. diversity) (6.3). The increase in the density of academic vocabulary over time broadly supports Storch and Tapper's (2009) finding reporting a growth in academic vocabulary tokens (measured on the basis of the AWL) in international postgraduate student writing over a period of ten weeks as well as Xudong et al.'s (2010) result showing a slight increase in the use of AWL tokens in international graduate students' writing over a similar period of time. However, this study's finding is contrary to other studies, which have found no changes in academic vocabulary tokens over ten weeks (Storch, 2009), one year (Knoch et al., 2014) and three years (Knoch et al., 2015) (3.2.5). While Storch (2009) attributed a lack of improvement in the usage of academic vocabulary in students' writing to a relatively short period of study (i.e. 12 weeks), this study has shown that the period of 24 teaching weeks over 6 months (i.e. the duration of the IFP) is sufficient for novice student writers to develop their productive knowledge of academic vocabulary.

In terms of the increased diversity of academic vocabulary over the duration of the IFP, comparison with other studies is difficult as very little was found in the literature in relation to the changes in the diversity of academic vocabulary in student writing. This is because most studies into the deployment of academic vocabulary focus primarily on the proportion (i.e. density) of academic vocabulary in student written production (3.2.5). An exception is Storch and Tapper's (2009) study reporting an increase in academic vocabulary types (measured on the basis of the AWL) in international postgraduate student writing over a period of ten weeks. Although their study was conducted over a shorter period and investigated writing produced by students at a higher level of academic study, the result of the current study is somewhat comparable with that of Storch and Tapper's (2009) and suggests that a period of six months (i.e. the duration of the IFP) is sufficient for foundation-level novice student writers to improve the repertoire of their academic vocabulary and demonstrate this improvement in written production.

6.4.2 The role of the topic and assignment brief

The findings generated in response to RQ2 confirmed some of the findings generated in response to RQ1 (Chapter 5); in particular, the impact of the topic on the deployment of academic vocabulary (discussed in 5.3.2), suggesting that some topics may prompt a denser and more diverse usage of academic vocabulary. The impact of the topic was also noted by Olinghouse and Wilson (2013), for instance, whose findings showed that topic knowledge is one of the knowledge bases accessed by students in order to select appropriate vocabulary during the process of composing written text (3.2.5) and is also corroborated by Knoch et al., (2015, 2014), whose studies spanned one (Knoch et al., 2014) and three years (Knoch et al., 2015) and showed no increase in the deployment of academic vocabulary. They theorise that the lack of improvement in the deployment of academic vocabulary in their studies could be explained in relation to the assigned topic, which may have allowed for a limited range of vocabulary (3.2.5).

In addition to the impact of the topic, the increase in the density of academic vocabulary was found to have been impacted by a denser usage of academic vocabulary drawn from the assignment brief, which also matches the findings relating to RQ1 (discussed in 5.3.1). Contrary to this, the increase in the diversity of academic vocabulary was not found to have been impacted by the academic vocabulary types contained in the assignment brief. Thus, in terms of developing the repertoire of academic vocabulary (i.e. diversity), the decrease in the diversity of academic vocabulary contained in the assignment brief suggests that the students drew on other sources towards the end of the IFP. This was confirmed by the higher number of reading sources used in E2, discussed next.

6.4.3 The role of reading sources

The use of reading sources was found to be one of the contributing factors that impacted the students' usage of academic vocabulary in their written assignments (further discussed in Chapter 7) as the number of sources on which the students drew in both E1 and E2 had doubled in E2 (from an average of 7 sources in E1 to 14 sources in E2) (6.3.3). The important role that reading sources played was further supported by the survey and interview data. This result is, however, contrary to Storch's (2009) finding showing little change in the students' use of sources over a 12-week period, but is in line with Storch and Tapper's (2009) study, in which they partially attributed

the increased usage of academic vocabulary tokens and types in the students' writing to exposure to academic texts relating to their studies.

This finding is also consistent with results generated by studies into the potential links between reading and writing (referred to as research into writing from sources, textual borrowing, source-based writing, reading-to-write constructs or integrated writing tasks) (3.2.2), which emphasise that the skill of writing is not autonomous, but rather allied with other skills such as reading, whereby the texts serve as a source of vocabulary for writers. Plakans and Gebril (2012), for instance, found that one of the uses of texts by writers was reading for vocabulary to support language in writing. Their finding corroborates that of Leki and Carson (1997), whose study showed that source texts provided students with vocabulary and writing style and thereby "often became scaffolding for the subsequent assignment by freeing the writer from the need to find appropriate words" (p. 56) (3.2.2).

Despite the vital role that reading sources play in the development of students' academic vocabulary, some researchers emphasise the importance of instruction which ought to focus on expanding L2 learners' lexical repertoire as "mere exposure to academic text and reading may not be sufficient for L2 learners to attain the advanced academic proficiency essential for success in their academic endeavors" (Hinkel, 2003, p. 297). This is also in accord with Storch and Tapper's (2009) study (3.2.5), in which they point to explicit teaching as a potential contributor to the increase of academic vocabulary items in their students' texts over time as students' attention was focused on appropriate expressions (including those contained in the AWL) for various textual functions, such as referring to sources, making comparisons, or commenting on research methods and results. Explicit focus on academic vocabulary in classes was also highlighted in Brun-Mercer and Zimmerman's (2015) study, for example, which found that the participants had difficulties deploying academic vocabulary effectively and appropriately as they were not always aware of the register of a word. They thus recommend explicit teaching of the register of new vocabulary. Csomay and Prades (2018) also underlined the importance of teaching academic vocabulary with a contextualised approach. Similarly, Cons (2012) calls for explicit instruction on effective integration of academic words in writing. Thus, although reading sources may play a vital role in the deployment of students'

academic vocabulary, reading alone may not be sufficient for successful integration of appropriate academic vocabulary items in students' written production.

6.4.4 Core academic vocabulary

Further changes were identified in the size of the core academic vocabulary, which showed an increase in the number of academic vocabulary types shared among the students. This increase was achieved by academic vocabulary types displaying certain textual functions. The highest increase was found in research-oriented procedural academic vocabulary, primarily represented by action words (4.4.2). These were mainly verbs in E1 (e.g. *apply*, *develop*, *improve*, *perform*) (Figure 6.9 in 6.3.5) with a higher number of nouns in E2 (e.g. *communication*, *interaction*, *performance, usage*) (Figure 6.10 in 6.3.5). This increase in the usage of nouns may partly be attributed to the students developing the skills of using a nominal style of writing, which is regarded as an important aspect of an academic writing style (Biber & Gray, 2016) due to the various functions that nominalisations have in academic texts, such as maintaining an impersonal tone or contributing to textual cohesion (Baratta, 2010) (2.2.3). These functions are considered relevant in the production of quality academic texts. Hence, this change in the students' deployment of academic vocabulary can be seen as important in the development of their academic writing skills.

The second-highest increase was found in research-oriented descriptive academic vocabulary, which showed a growth in evaluative adjectives in E2 (e.g. *important*, *negative*, *beneficial*) (Figure 6.12 in 6.3.5) compared to E1. This indicates that the students began to incorporate evaluative expressions when presenting alternative positions as required by this writing genre (i.e. a Discussion Essay), whereby fulfilling this genre's requirement to credibly argue for a position (Table 4.6 in 4.3.4). Considering the persuasive intent of argumentation required by this writing genre, such evaluative expressions are regarded as playing an important role in signalling the author's position as evaluative language is considered an important element contributing to the quality of written argumentation (Mei & Allison, 2005). Further, the importance of integrating evaluative language in this writing genre has important implications not only for argument structure, but also for the extent to which writers use evaluative wording effectively to express judgments and provide reasons for them together with supporting evidence (Mei & Allison, 2005). The increased usage of these research-oriented descriptive academic vocabulary items displaying an

evaluative function can thus be seen as an important step in the development of the student argumentative writing, particularly since this aspect of academic writing is considered to "pose a continuing challenge for learners" (Mei & Allison, 2005, p. 107). Therefore, similar to the findings generated in response to RQ1 (Chapter 5), the effect of the writing genre becomes evident.

A slight increase was also found in text-oriented framing academic vocabulary items (e.g. *include*, *both*) (Table 6.6 in 6.3.5), which were found to be used to introduce alternative positions on the issue under discussion as well as to provide relevant examples supporting the claims. Further, an examination of the text-oriented transition markers showed that although there was no overall increase in academic vocabulary items displaying this function (Table 6.6. in 6.3.5), there were differences in the academic vocabulary types deployed in the Essays. These were found in inclusion of academic vocabulary items fulfilling the function of expressing a contrast (e.g. *however*) in E2 (Discussion Essay), compared to markers used to indicate addition only (e.g. *furthermore*, *moreover*) in E1 (Exposition Essay). This shows that the students were aware of both genres' requirement; that is, the Exposition Essay did not require the students to present alternative arguments and hence the students did not have an opportunity to integrate transition markers expressing a contrast. The inclusion of transition markers in the Discussion Essay could thus point to the students' awareness of the writing genre's requirements to present and discuss alternative positions before reaching a position.

The text-oriented framing and transition academic vocabulary items correspond to the textual category of metadiscourse (4.4.3); namely to the sub-categories of 'frame markers' signalling boundaries in the discourse or stages in the argument, and 'transition markers' expressing semantic relation between clauses (Hyland, 2005). This integration of metadiscourse is important in the development of student writing due to its function as a link between a text and its context primarily because readers have certain expectations of the organisation of the text as well as of the inclusion of sufficient signals of the writer's intentions and views (Hyland, 2005). Further, the awareness of metadiscourse features offers several advantages to student writers, such as helping them to "better understand the cognitive demands that texts make on readers and the ways writers can assist them to process information" (Hyland, 2005, p. 178). In addition, metadiscourse provides student writers with "the

resources to express a stance towards their statements" and allows them to "negotiate this stance and engage in a community-appropriate dialogue with readers" (Hyland, 2005, p. 178).

Hence, due to the important role that metadiscourse plays in texts, it is regarded as an essential element of argumentative discourse as it shows "how writers seek to influence readers' understandings of both the text and their attitude towards its content and the audience" (Hyland, 1998, p. 437). In this study, therefore, the appropriate integration of metadiscourse in both Essays can be seen as an important feature in the students' assessed academic writing as the ability to use metadiscourse effectively is considered a defining element of successful academic writing (Hyland, 2005).

6.5 Development of academic vocabulary: summary

This chapter presented findings and discussion in response to RQ2 (i.e. To what extent do the density and diversity of academic vocabulary in international foundation students' assessed academic writing develop over an academic year?). The investigation into the development of academic vocabulary in the students' assessed academic writing over the IFP, measured in terms of an increase in the density (6.2) and diversity (6.3) of academic vocabulary between the first (i.e. Exposition Essay, E1) and the last (i.e. Discussion Essay, E2) written assignment, showed an overall increase in both the density and diversity of academic vocabulary. That is, this improvement was found in the increased coverage or proportion of the students' written assignments formed by academic vocabulary items (i.e. the density of academic vocabulary) as well as in their repertoire of productive academic vocabulary (i.e. the diversity of academic vocabulary).

In both instances, this increase was achieved by a large proportion of students with 77% of students demonstrating an increase in the density of academic vocabulary and 70% of students showing an increased diversity of academic vocabulary. In addition, the increase in both the density and diversity of academic vocabulary was caused by two out of the three sub-corpora only, suggesting that some topics may have prompted a denser and more diverse usage of academic vocabulary than others. Next, unlike the increase in the diversity of academic vocabulary, which seemed to have resulted from drawing on a wider use of reading sources, the increase in the

density of academic vocabulary was found to be impacted by higher density of academic vocabulary items drawn from the assignment brief, suggesting that the students developed the strategy of addressing the assignment instructions by explicitly referring to the academic vocabulary items contained in the assignment brief, which led to a denser usage of these academic vocabulary items in their final written assignment. Further, the size of the core academic vocabulary increased and changes in the functions of the core academic vocabulary were also identified. These changes were found to have been impacted by the writing genre's social purposes, with another important textual feature arising from the genre's social purposes closely relating to metadiscourse. In addition, the students were found to have developed a nominal style of writing.

To complement these findings based primarily on textual analyses, the next chapter considers the students' perceptions of the main contributing factors that impacted the acquisition, deployment and development of academic vocabulary in their written assignments on the IFP (RQ3).

Chapter 7: Students' perceptions

7.1 Introduction

This chapter presents findings together with a discussion relating to the IFP students' perceptions of the main contributing factors that impacted their productive academic vocabulary used in their assessed academic writing (RQ3). For the purpose of exploring the students' perspectives, online questionnaires from the overseas participants (N=51) and interview data (N=14) from students across the three campuses (comprising 4 participants from the UK-based campus and 5 students from each of the two overseas campuses) were collected (4.3.5) and analysed (4.4.4). From these data, three aspects have been identified relating to the students' experience with academic vocabulary on the IFP: the acquisition (i.e. the sources) of new academic vocabulary based on the insights gained from the interview data (7.2); the deployment (i.e. the productive usage) of academic vocabulary informed by the questionnaires (7.3); and the development (i.e. improvement in the productive usage) of their academic vocabulary (7.4) identified on the basis of the interview data. This chapter concludes with a summary of the main findings (7.5).

It is noteworthy that although the collection of the interview data was motivated by the academic literacies approach to exploring student writing from the writers' perspectives, often utilising 'talk around text' (2.3), it was not possible to employ this approach to gaining insights into the participants' writing practices in the present study. This was due to the timing of the collection of the textual and interview data; specifically, since this study was conducted independently from the teaching on the IFP, both textual and interview data were collected after the academic year (4.3.4) with the interviews taking place before the textual analysis was completed. This meant that it was not possible to discuss the students' writing practices as they engaged in the process of completing their written assignments during the academic year, nor was it possible to discuss specific areas of interest generated by the textual analysis. The interview data were thus used to enable the students to reflect on their experience with academic writing during the IFP.

7.2 Acquisition of academic vocabulary

From the 14 interviewed students, two main themes emerged as the main contributing factors which were classified as relating to the acquisition of new academic vocabulary: appropriate sources (7.2.1.) and instructed environment (7.2.2). These themes were identified on the basis of responses to question 8 (below), mainly to the second part of the question relating to the perceived contributing factors.

8) Vocabulary development

Q: (How) do you think your academic vocabulary used in written assignments has developed over the course of the academic year?

Q: What do you think has contributed to this?

(prompts, if needed: taught sessions, independent learning, reading)

7.2.1 Appropriate sources

The use of or exposure to appropriate sources was mentioned by the majority of students (11 out of 14 students; 79%) during the interviews. This theme was found to encompass the sub-themes of reading and reference sources.

Reading sources

The perceived importance of reading sources is illustrated by the following quotes:

> "The readings and the researching I did help me a lot to pick up certain words … there are like special words that I picked up from the other writers" (Student UK16/17/A).
>
> "By reading new books and by reading academic articles …; by reading academic sources I got the vocabulary from there." (Student UAE18/B).
>
> "[New words came from] the materials you are reading for your coursework because you have to do research so sometimes you get to learn new words when you come across a word you don't know in a material and if you want to use the material you have to check it out and I think it helps" (Student MRU17/18/O).

These interview excerpts show that the students were aware of the key role that appropriate reading sources played in the acquisition of academic vocabulary. This

finding is supported by research on vocabulary learning through reading in both L1 and L2 contexts and is related to implicit or incidental learning, which often involves inferring the meaning of words from context as a primary vocabulary skill (3.2.2). Research on L1 reading has shown that reading and vocabulary growth are closely related, emphasising the need for multiple opportunities to read. Carver (1994), for instance, suggests that if learners are to expand their vocabulary, they need to be exposed to reading materials that are not too easy for them, otherwise they will not encounter many unfamiliar words. Research on learning L2 vocabulary through reading also shows that reading can lead to the learning of small amounts of vocabulary; in particular, texts which contain repetition of unfamiliar vocabulary (i.e. texts on the same topic) provide favourable conditions for vocabulary learning (Nation, 2001). This also reflects a determination discovery vocabulary learning strategy, which involves guessing words' meanings from contexts (Schmitt, 1997) (3.2.2).

This study's finding relating to the importance of reading sources for academic vocabulary acquisition is also in line with other studies conducted in the context of academic vocabulary research (3.2.2 & 3.2.5). Nagy and Townsend (2012), for example, see reading texts as particularly important for the learning of general academic vocabulary, supported by Nation (2001) who notes that meeting vocabulary in texts is the initial opportunity to meet new words. The vital role of reading sources is also emphasised by Zimmerman (1997), who found that students who tended to read the least reported the most difficulty with academic words. She thus emphasises the benefits of reading in acquiring new vocabulary and considers it one way for students to gain lexical knowledge as reading texts present words in meaningful contexts. This is valued highly by students since such contexts provide information about the properties of a word, which contributes to the learners' knowledge about the multifaceted nature of words (Zimmerman, 1997). Krashen (2012) also highlights the importance of reading and suggests that it plays a more important role in the learning of academic vocabulary than instruction: "there is good reason to hypothesize that academic vocabulary is acquired gradually through genuine academic reading for the readers' own purposes and that this path is more effective and efficient than even rich instruction" (p. 233). This is corroborated by Nagy (1997, p. 75), who believes that reading can be regarded as more powerful than

teaching as it is “unlikely that instruction accounts for anywhere near as much vocabulary growth as does incidental acquisition from context during reading". Storch and Tapper (2009) also partially attribute the improvement in students’ academic vocabulary (measured on the basis of the AWL) to exposure to academic texts.

Hence, the IFP students’ awareness of the role of reading source can be seen as an important element in the acquisition of their academic vocabulary forming an indispensable element of academic language necessary in university contexts. As noted previously (6.4.3), however, some researchers emphasise that reading alone may be insufficient for L2 learners to acquire the necessary academic proficiency and emphasise the importance of explicit teaching of academic vocabulary (Brun-Mercer & Zimmerman, 2015; Cons, 2012; Csomay & Prades, 2018; Hinkel, 2003; Storch & Tapper, 2009). This study’s findings also suggest that in addition to reading sources there were other factors that played an important role in the acquisition of academic vocabulary (discussed below).

Reference sources

In addition to reading sources, two students noted the thesaurus as a source of new vocabulary:

> “Thesaurus helps a lot because you get to get a lot of different words that you can use to substitute the other words” (Student MRU17/18/O).

> “[new words came from] thesaurus when I’m doing paraphrasing or summarising just to give it my own words from there” (Student MRU18/E).

This highlights the important role that suitable reference materials play in the acquisition of academic vocabulary, particularly monolingual reference sources (i.e. thesauruses) offering a range of synonyms. This finding is in line with Salehi and Habibi's (2015) study reporting exposure to numerous different words with the same meaning among the benefits of using a thesaurus (3.2.2).

Another source of acquisition of academic vocabulary, mentioned by one student, was the Academic Phrasebank - an online resource containing examples of mostly content neutral (i.e. discipline non-specific) academic phraseological expressions designed primarily for L2 academic writers (The University of Manchester, 2020):

> "During the IFP we used the Academic Phrasebank; there were a lot of words to use in essays and reports, so I picked them up from there" (Student UAE17/18/N).

Drawing on the Academic Phrasebank (containing phrases as opposed to individual vocabulary items) shows the Student's awareness of the importance of individual words used in their phraseological environments in written production, also found by Staples et al. (2013) for instance, who emphasise that phraseologies are "an essential part of native and native-like language use" (p. 214). The importance of phraseologies in academic contexts has also been highlighted by others (e.g., Hyland, 2008; Martinez & Schmitt, 2012; Simpson-Vlach & Ellis, 2010) (3.2.4) as the absence of phraseologies may indicate lack of expertise in academic context (Qin, 2014, p. 220).

The use of reference materials was thus found to be another means of vocabulary acquisition for the participants in this study, which is also one of the determination discovery vocabulary learning strategies (Schmitt, 1997) (3.2.2).

7.2.2 Instructed environment

A further recurrent theme in the interviews was the importance of instructed environment (i.e. teaching, lectures, seminars), reported by over a half of the interviewed students (8 out of 14 students; 57%) as one of the perceived sources of the acquisition of academic vocabulary, illustrated by the interview excerpts below.

> "During classes…the teachers when they explain to us…we are trained to focus on academic [vocabulary]" (Student UK16/17/A).

> "The teaching definitely helped … I think the teaching was the biggest thing" (Student UK16/17/B).

> "In our sessions there were introducing us new topics and with those new topics we learned new words too" (Student UAE17/18/UU).

This study's finding relating to the role that explicit teaching of academic vocabulary plays is consistent with other research emphasising the importance of instructed environment in the context of academic vocabulary learning. Storch and Tapper (2009), for instance, suggest that their participants' improvement in the usage of academic vocabulary over time may have resulted from attention being drawn to appropriate expressions (including those

appearing in the AWL) on the EAP course that the participants completed (3.2.5). The importance of conscious study of new words is also emphasised by Nation (2001), who notes that although learners may initially meet new words in texts, this meeting needs to be combined with intensive study (3.2.2). Similarly, Corson (1997) points out that the teacher is indispensable in formal education as the most basic kind of assistance that academic learning can receive is provided by the teacher, which is considered useful for initial access to academic meanings. This is supported by Cons (2012), who also emphasises the importance of instructed environment in the process of acquiring academic vocabulary items (3.2.5).

Specifically, Schmitt and McCarthy (1997) propose that the most frequently occurring words should be given explicit attention as they consider it a good way to acquire some initial information about the newly met words. The benefit of such explicit learning, it is argued, lies in attention being focused directly to the information that the learners need to learn, thereby providing the greatest chance for acquisition (Schmitt, 2000) (3.2.2). This is further supported Carlo et al. (2004), who note that whichever words teachers select to teach, they need to give learners ample structured opportunities to encounter these new words in authentic and engaging contexts. Relating more specifically to the usage of academic vocabulary in written assignments, Nation (2001) suggests that teachers could pre-teach academic vocabulary that they consider necessary for assignments, for instance.

The key role of the instructed environment is encompassed by Cazden (1988 in Corson, 1997, p. 703) who sees the "difference between what people can do on their own in using language, and what they can do with the help of an older and more experienced language user" as important, closely relating to a sociocultural theory of learning; in particular to Vygotsky's notion of the 'zone of proximal development' (2.2.1), which has been influential in education and L2 teaching. This finding also corresponds to explicit or intentional vocabulary learning whereby instructed environment provides the students with conscious and focused study of words involving activities intended to result in vocabulary learning, and can also be seen as relating to the social discovery vocabulary learning strategy (Schmitt, 1997) (3.2.2).

7.2.3 Acquisition of academic vocabulary: summary

The above findings relating to the acquisition of academic vocabulary point to two areas which the IFP learners in this study perceived as the main contributing factors impacting the acquisition of new academic words: exposure to suitable reading and reference sources and through instruction. Since academic vocabulary is a type of vocabulary (3.2.1), the same general processes apply that may help learners acquire a new vocabulary item (Nation, 2001). These processes comprise noticing (which can occur through formal instruction or textual input), followed by retrieval and generative use of the newly acquired vocabulary item (3.2.2). Therefore, learners first need to notice the word and consider it a useful vocabulary item. Some of the instances when noticing occurs is when a new word is explained to them (e.g. explicit teaching of vocabulary items), due to the salience of the item in the textual input (e.g. the use of appropriate reading sources), or it can also occur when learners use a reference source to look up an item. These instances of noticing are in line with the above presented findings, which point to instructed environment and the use of suitable sources (e.g. reading texts or reference sources) as the primary factors facilitating the acquisition of academic vocabulary, also corresponding to discovery vocabulary learning strategies (3.2.2).

7.3 Deployment of academic vocabulary

In terms of the deployment of academic vocabulary (i.e. the productive usage of academic vocabulary items) in the students' written assignments, the survey data (N = 51) were the main source of information, primarily question 14 (below).

14. Which of the following writing strategies did you use when completing your essays and reports on the IFP? (tick as many as you like)

- paraphrasing (i.e. using own words)
- summarising (i.e. extracting main points)
- translating
- using direct quotes (i.e. using the wording of the source text)
- using vocabulary from texts
- using vocabulary from assignments briefs/instructions
- using vocabulary from teaching materials
- other…

From the survey data, the importance of various textual sources (7.3.1) became apparent.

7.3.1 Textual sources

The important role that various textual sources play in the process of completing written assignments became evident from the survey data, where approximately 43% of the students (22/51) reported using vocabulary from teaching materials as one of the strategies used when completing their written assignments. The assignment brief emerged as another source of vocabulary with nearly 40% of students (20/51) stating that they had drawn on the vocabulary contained in the assignment brief in their writing. Around 30% of students (16/51) reported using direct quotes and a similar proportion of students (15/51) noted using vocabulary from texts during the process of completing their written assignments. These results are further discussed next.

Teaching materials

Using vocabulary from teaching materials during the process of completing their written assignments was selected by the highest proportion of students (approximately 43%). This result, to some extent, supports findings relating to the acquisition of academic vocabulary through instructed environment (7.2.2), which showed the students' perceptions of teaching as an important factor contributing to the acquisition of academic vocabulary. The emphasis placed by students on teaching materials during the written production stage can thus be seen as a subset or by-product of explicit teaching of academic vocabulary items, which was reported to be important in the acquisition of academic vocabulary. This finding is consistent with that of Storch and Tapper (2009), whose study showed an improvement in the students' usage of academic vocabulary (i.e. AWL items) over time. They partially attribute this improvement to the EAP course that the students attended, which provided the learners with materials on academic language use. These materials contained model texts, for instance, which were analysed for structure and language, or included various language exercises on linguistic choices appropriate for academic register, followed by opportunities for language production. Storch and Tapper (2009) thus speculate that use of appropriate teaching materials was one potential factor contributing to the improvement of the students' academic vocabulary usage (3.2.5).

The assignment brief

As far as the assignment brief is concerned, noted by around 40% of students, this result shows the students' reliance on the academic vocabulary contained in the assignment brief, also supported by the findings generated by textual analyses in response to RQ1 (5.3.1) and RQ2 (6.2.2 & 6.3.2). These findings suggested that drawing on the lexical items contained in the assignment brief may have been a strategy employed particularly by less proficient students, who were less confident in using appropriate vocabulary and their own words (5.3.1), but it may also have been used as a strategy of addressing the assignment instructions (6.2.2 & 6.3.2). The students' usage of vocabulary items contained in the assignment brief was also reported by Flowerdew (2003), for instance, who traced some of the expressions used in the students' writing investigated back to the assignment instructions. Similarly, Milton (2001) also reported in his study of Hong Kong Chinese university student writing that some students used the vocabulary from the assignments prompts (2.2.3 & 3.2.2).

Direct quotes

The use of quotation, reported by approximately 30% of the 51 students who completed the questionnaire, is in line with Coxhead's (2012) study, which found that incorporating quotation from a source text was used by some writers in her study as a way to include specific words in their writing. For some of these writers, using quotation felt like a safe way of using academic words. This rationale for using quotations was also noted by Student MRU18/E, who stated in the questionnaire that s/he was relying on using direct quotes first as s/he was not confident in paraphrasing:

> "When I started I was just doing quotes because I found paraphrasing to be a bit difficult. Now I think I've got the hang of paraphrasing, so I do a lot more paraphrasing" (Student MRU18/E).

It is thus reasonable to assume that for less confident writers, the use of direct quotes may be a valuable strategy that enables them to incorporate unfamiliar academic vocabulary items in their written production. This can be seen as a useful strategy since an initial reliance on quotation may particularly help novice academic writers to familiarise themselves with newly met academic vocabulary items and assist them

in engaging with both the meaning and form of these words, which may lead to increased confidence in deploying these vocabulary items in written production.

Reading texts

The use of quotation can be seen as related to the reliance on reading materials during the writing process, as reported by approximately 30% of students. This result corroborates the findings generated in response to RQ2 (Chapters 6), which showed the impact of reading sources on the diversity of academic vocabulary, whereby drawing on a higher number of reading sources resulted in greater variety of academic vocabulary (6.3.3). Leki and Carson (1997, p. 41) define such sources as texts in the broadest sense of term "to which writers are exposed and required to account for in some way". Written production based on textual sources requires the writers to demonstrate an understanding of the source texts; that is, writers must produce "text-responsible prose based on content acquired primarily from text" (Leki & Carson, 1997, p. 41).

This finding thus supports that of Leki and Carson (1997), which showed that source texts were often utilised in assignments by providing the writers with appropriate vocabulary. In particular, they note the advantages of writing with source texts as well as the disadvantages of writing without a textual source, as reported by the university student participants in their study. Their findings show that the writers appreciated writing with a textual source as it could supply numerous resources, such as vocabulary and writing style. As a result, the writers in their study perceived writing with a source text as easier than writing with no textual source (Leki & Carson, 1997, p. 51) (3.2.2). Conversely, some of the disadvantages of producing a piece of writing without a textual source were found to revolve around deficits, including lack of familiarity with vocabulary relating to the topic (Leki & Carson, 1997). They also note that in their study, texts were found to be used as "scaffolding for the subsequent assignment by freeing the writer from the need to find appropriate words or to figure out the appropriate rhetorical form" (Leki & Carson, 1997, p. 56).

This study's finding relating to the benefits of drawing on sources is also in line with research into source-based writing (also known as textual borrowing, reading into writing, writing from sources, reading-to-write constructs or integrated writing, discussed in 3.2.2) investigating the use of a source text during the writing process, particularly the integration of vocabulary items in written production, where several

researchers have noted the important role that various textual sources play in the process of completing written assignments from a vocabulary perspective. Plakans and Gebril's (2012) study, for example, reports that one of the uses of texts by writers was reading for vocabulary to support language in writing (3.2.2). The importance of exposure to academic texts during the writing process is also emphasised by Cons (2012, p. 630): “Academic texts should be provided to students to use as models for their own writing. If students can follow an example text, this can help them feel more comfortable using academic words in writing” (3.2.5). This study’s finding relating to the importance of a textual input during the writing process is thus in line with previous research on writing from sources.

7.3.2 Deployment of academic vocabulary: summary

The above presented findings relating to the deployment of academic vocabulary in the students’ written production identified various textual sources as a contributing factor that impacted the students’ usage of academic vocabulary in their written assignments. Among these textual sources were the teaching materials, the assignment brief and reading texts that the students drew on during the writing process, with the reading text having provided them with quotations subsequently incorporated in their writing.

These findings can be seen as being in line with all three stages of learning new vocabulary (i.e. noticing, retrieval and generation) (3.2.2) as follows: the textual sources can enable noticing of new academic vocabulary items as well as the retrieval of previously noticed academic vocabulary, followed by the generation stage, whereby the students deploy the vocabulary items they had met in the various sources in their written assignments in ways that are different from the previous encounters with the words.

7.4 Development of academic vocabulary

Insights into the students’ perceptions of their academic vocabulary development was gained from the interview data, where question 8 (below) was the main source of information.

8) Vocabulary development

Q: (How) do you think your academic vocabulary used in written assignments has developed over the course of the academic year?

Q: What do you think has contributed to this?

(prompts, if needed: taught sessions, independent learning, reading)

Development (i.e. perceived improvement in the productive usage) of academic vocabulary was reported by the vast majority of the interviewed students; that is, eleven of the fourteen (79%) students said they thought their academic vocabulary had improved during the course of the IFP:

> "I started to write good vocabulary, I learned a lot of new words, that was probably the biggest achievement I did" (Student UAE18/B).

> "I think my vocabulary has widened" (Student MRU18/E).

Two students were unsure whether their academic vocabulary had improved:

> "…not sure I have improved [academic vocabulary] or not…I think now I have much more vocabulary compared to the beginning [of the IFP]" (Student UK16/17/C).

> "I don't know, I'll wait till somebody tells me that because I'm still working on myself" (Student MRU18/F).

From the interviews, several themes were identified as perceived contributing factors that impacted the development of the students' productive academic vocabulary in response to question 8 (above). These are: feedback (7.4.1), practice (7.4.2) and peers (7.4.3), discussed below.

7.4.1 Feedback

Feedback was highlighted by eight students (57%) in the interviews as one of the main factors that helped them improve their academic vocabulary in their written assignments:

> "I think having the opportunity to talk and connect with my tutors… because I needed their feedback before handing it in; I think that helped a lot" (Student UK16/17/B).

> "We can send drafts to the teachers and I can remake it from the feedback so it's not something that I made all by myself" (Student UK16/17/C).

> "I also had the room of submitting drafts before submitting your piece of work so in case there is any corrections I get feedback from the lecture so that really helps me" (Student MRU18/F).

All the above quoted students emphasised the importance of formative feedback; that is, feedback on drafts or work in progress, which is not marked. One student also reported the usefulness of summative feedback, referring to post-submission feedback which includes a mark:

> "The feedback was very helpful, that was the most helpful part of it, the feedback you get after each coursework was very helpful in this process" (Student MRU17/18/O).

The interview data are supported by the survey data, where receiving feedback on drafts was stated by 71% (36/51) of students as particularly helpful in the process of completing written assignments on the IFP. This finding hence shows that students generally appreciate receiving feedback on their work, particularly prior to submitting an assignment.

The importance of feedback accords with other studies investigating academic vocabulary (3.2.5). Knoch et al. (2014), for instance, report in their study that "students appreciate feedback on their writing and if this is lacking, are then not aware of any deficiencies in the quality of their texts" (p. 13). They also emphasise that generic feedback on writing, which fails to draw the learners' attention to language use, "may be a disservice to students" (p.12) as it may leave them with the impression that their writing does not need to be improved. The absence of feedback was also noted in Knoch et al.'s (2015) study, in which they concluded that lack of improvement in L2 writing, including lack of improvement in academic vocabulary, could be attributed to lack of feedback on the students' writing.

The importance of providing students with feedback on their writing in general, not only in relation the academic vocabulary usage, has been emphasised by many. Ferris (2003), for example, argues that the provision of feedback on student writing is the most important factor in the learners' writing development. Wingate (2010) also

found that formative feedback is effective in improving student academic writing. Weaver (2006), who investigated students' responses to tutors' written feedback, found four main feedback areas that students considered unhelpful. These include comments that were too general or vague, lacked guidance, focused on the negative or were not related to the assessment criteria. Other reported reasons why students fail to consider tutor feedback is their lack of understanding of the tutors' comments or because the feedback given does not motivate or guide them sufficiently. Weaver (2006) thus suggests that practitioners should ensure that their feedback is not only constructive and clear, but it is also related to assessment criteria and gives guidance on how to improve future performance.

This finding also relates to the importance of dialogic feedback, which focuses on the students' text in process as opposed to providing feedback on a written text as a completed product. This approach to feedback is associated with sociocultural traditions and is also closely related to the reconceptualisation of the widespread practice of 'feedback' as 'talkback' prominent in the academic literacies approach (2.3). The perceived importance of this dialogic approach was noted by Student UK16/17/B, for example, who found "having the opportunity to talk and connect with my tutors" helpful in the process of completing their written assignments. This can also be seen as connected to the academic literacies principle of foregrounding practice over text, in line with the view of academic writing as a social practice (2.3).

As far as summative feedback is concerned, noted by one participant in this study, some studies report that post-submission feedback has less value. Leki (2006), for example, points out that such feedback is often ignored by students, and in some cases not even seen if the assignment was submitted at the end of the term. In such instances, Chang (2014) argues that once the students submit their assignments, the learning process comes to an end as the feedback will not lead to further learning or improvement in academic writing.

From the above findings and discussion, the value of formative feedback emerges as an indispensable aspect of student writing development in general, and academic vocabulary in particular.

7.4.2 Practice

Being provided with sufficient opportunities for practice was mentioned by approximately a third (5 out of 14; 36%) of students as a perceived factor that had contributed to the improvement of their productive academic vocabulary. This included being given the opportunity to work on formative tasks:

> "We had a lot of formative tasks as well that we had to do" (Student UK16/17/A).

> "It's nice to have the formative, to have a chance to write something that you don't feel the pressure of it being graded" (Student UK16/17/B).

> "[Vocabulary improved owing to] writing all those papers, we had to use different words, keep the paper not in the same routine and things like that so we had to use different words so definitely it [academic vocabulary] did [improve]" (Student UAE17/18/UU).

The importance of providing students with writing practice has also been highlighted by several vocabulary and writing researchers (3.2.5). Cons (2012), for instance, suggests that writing ought to be assigned frequently as English learners "need more writing practice in general to become more comfortable with the act of writing so that they will feel more comfortable using more words overall, specifically academic words" (p. 630). This is supported by Knoch et al. (2015), who also found that the participants in their study who did not think their writing improved attributed this primarily to insufficient writing practice. This accords with Nagy and Townsend (2012, p. 96), who note the value of opportunities to practise newly acquired academic vocabulary items in authentic contexts in order to enable students to use them in the relevant contexts. Similarly, Carlo et al. (2004) suggest that providing opportunities for learners to practise newly acquired vocabulary items through writing is crucial. Coxhead (2012) also calls for practitioner to consider the amount of academic writing practice that learners receive. Similarly, Neumann (2014, p. 92) points out that "students need a space without marks to really improve and expand their language repertoire" as they may choose to follow the familiar in assessment situations and not take risks. Such risk avoidance may be seen as detrimental to the students' language development since risk-taking in language production has been found to be associated with L2 learning success (Neumann, 2014).

7.4.3 Peers

Three students (21%) perceived peer support as one of the contributors to the improvement of their academic vocabulary:

> "[I discussed] the different style of writing with my friends, I took the essays of those who got higher marks and read it and see how I can improve on that… I had my friends check my work for the mistakes so [for the last essay] we wrote as a group sharing ideas, the content, helping each other so it was a much better result in the end" (Student MRU18/C).
>
> "The group work you have, the group activities, that helps because you get to benefit from each other" (Student MRU17/18/O).
>
> "There were students from various countries so the only language we could speak was English so that I guess that also helped me improve my vocabulary by learning new words from my friends and my classmates" (UAE18/B).

The three students who pointed out the benefits of group work highlight the importance of peer support and the opportunity to learn from peers. The important role that peers can play in language development is also noted by Corson (1997), who suggests that even though the teacher often provides initial access to academic meanings, it is later dialogue that provides the necessary elaboration. This dialogue can be with classmates, for instance, and its benefit lies in providing "important series of reconceptualizations needed to master rules of use across different contexts" (Corson, 1997, p. 703). Group work or a similar dialogic activity thus needs to follow the initial conceptualization in order to benefit the language output (Corson, 1997).

This study's finding relating to the benefits of learning vocabulary from peers in group work corroborates other vocabulary researchers (3.2.2). Huong (2006), for example, found that the students in his study were drawing on each other's vocabulary knowledge by either asking other group members when they met a new word, or they learned new words by listening to the interaction among the group members. Huong's (2006) study thus highlights the advantages of group work as it shows that working in groups, particularly if the groups include a more capable peer, gives the group members an access to someone who may know the required new vocabulary items which convey the concepts and ideas necessary for the group discussions.

Dobao's (2014) study also showed that group work is beneficial to vocabulary learning as interaction in groups leads to significantly more instances of second language vocabulary learning than pair interaction. Similarly, Lin's (2018) study also reported considerably higher vocabulary gains when working in groups compared to working individually. Specifically, some of the advantages of group work due to its interactive nature reported in Lin's (2018) study include receiving and giving information, discovering richer lexical information as well as richer support for retention.

Considering the benefits of group work or peer support in relation to vocabulary development, some researchers have called for the use of collaborative written tasks (e.g., Jelodar & Farvardin, 2019; Swain, 2001) as they believe that by working on a common goal and sharing the responsibility for the final writing product, group members are encouraged to discuss the language they are using and draw on each other's linguistic resources. In doing so, even learners at the same level can offer scaffolded assistance to other group members as no two learners possess the same strengths, weaknesses, knowledge or resources, whereby they can achieve a level of performance beyond their individual level of competence (Dobao, 2014).

This study's finding relating to the IFP students' perceived benefits of peer support, often in the form of group work with emphasis on assistance of a more capable peer, is in line with a sociocultural theory of learning, particularly Vygotsky's notion of the 'zone of proximal development' (2.2.1). It is also one of the social discovery as well as consolidation vocabulary learning strategies (Schmitt, 1997) (3.2.2).

7.4.4 Development of academic vocabulary: summary

The above results highlight three areas that were reported to impact the improvement in the productive usage of academic vocabulary by the IFP students in this study. These include the key role of receiving feedback on academic written production in general and on the usage of academic vocabulary in particular, the importance of being provided with opportunities to practise the deployment of academic vocabulary in writing and the value of peer support.

Two of these three reported factors, namely peer support and practice, also match Nation's (2001) processes of vocabulary learning (3.2.2). Through peer support or group work, all three processes involved in the learning of new words (i.e. noticing,

retrieval, and generation) are fulfilled as follows: during the interaction enabled by group work, learners may notice new vocabulary items during discussions with their peers, retrieve them both receptively (i.e. when the word is met in listening or reading) or productively (i.e. retrieving the spoken or written form of the word), and follow it by generative use (i.e. previously met vocabulary items are used in new ways) in collaborative writing tasks. The finding relating to the benefits of practice relate primarily to the third stage of vocabulary learning (i.e. the productive generative use of newly met academic vocabulary). In addition, the perceived benefits of feedback and peer support can be seen as closely relating to Vygotsky's notion of the 'zone of proximal development' (2.2.1), influential in education as well as in L2 teaching contexts with peer support also corresponding to social discovery and consolidation vocabulary learning strategies (3.2.2).

7.5 Students' perceptions: summary

This chapter presented findings and discussion in response to RQ3 (i.e. What are the students' perceptions of the contributing factors impacting the acquisition, deployment and development of their academic vocabulary?). From the interview and survey data, exposure to appropriate reading and reference sources together with instructed environment were found to play an important role in the acquisition of academic vocabulary, while other textual sources (including teaching materials, the assignment brief and reading texts) were reported to impact the deployment of academic vocabulary in the students' written assignments. Feedback, peers and opportunities for practice were identified as perceived factors impacting the development of academic vocabulary.

The perceived factors reported by the IFP students' as having impacted the acquisition of their academic vocabulary also represent both implicit (or incidental) vocabulary learning (i.e. exposure to appropriate reading and reference sources) as well as explicit (or intentional) vocabulary learning (i.e. instructed environment). Some of the factors perceived as having contributed to the IFP students' vocabulary learning also relate to Schmitt's (1997) discovery vocabulary learning strategies (i.e. the use of appropriate sources, instructed environment and peer support) and consolidation vocabulary learning strategies (i.e. peer support).

Chapter 8: Conclusion

This study set out to investigate the usage of academic vocabulary in international foundation-level students' assessed academic writing at a UK university. The motivation behind the study was multi-fold. The primary driver for the research was my experience of working with multilingual foundation-level students in the context of UK tertiary education and my personal interest in university students' academic writing. The rationale for this study also stemmed from the importance of foundation-level provision within the UK higher education climate, which warrants attention due to the crucial role that these courses play in preparing students for undergraduate study (1.3). Next, considering that written assignments are the main form of assessment at universities (1.2), it is important to understand novice students' academic writing so as to be better able to assist them in developing their academic writing skills. In particular, given that academic vocabulary is a key element of academic writing style (3.2.3), gaining an understanding of the learning and subsequent deployment of this specialised type of vocabulary in students' written production is crucial. Further, despite an extensive body of research into various aspects of university student writing (1.4), including the usage of academic vocabulary (3.2.5), general academic vocabulary in the academic writing of students at a foundation level of study remains under-researched (3.2.6). This study has thus attempted to address this important omission by examining the assessed academic writing of a multilingual group of foundation-level students in the context of a UK university with a particular focus on academic vocabulary due to the vital role that these vocabulary items have in EAP contexts and wider university settings.

This concluding chapter first provides a summary of the main findings (8.1), then focuses on the contributions that this study has made to research (8.2.1), methodology (8.2.2) and its implications for pedagogy (8.2.3). This is followed by the limitations of the current study (8.3), which lead to recommendations for future research (8.4).

8.1 Summary of findings

The overarching aim of the present study was to explore the usage of academic vocabulary in international foundation-level students' assessed academic writing. Specifically, this study focused on the deployment of academic vocabulary across four

writing genres of assessed academic writing (i.e. Exposition Essays, Problem Questions, Research Reports and Discussion Essays) produced by a multilingual group of foundation students, addressed by RQ1 (i.e. What are the density and diversity of academic vocabulary in international foundation students' assessed academic writing across writing genres?). It also investigated the changes in the deployment of these vocabulary items over the academic year in response to RQ2 (i.e. To what extent do the density and diversity of academic vocabulary in international foundation students' assessed academic writing develop over an academic year?). In addition, the students' perceptions of the main contributing factors impacting the acquisition, deployment and development of academic vocabulary were explored by RQ3 (i.e. What are the students' perceptions of the contributing factors impacting the acquisition, deployment and development of their academic vocabulary?).

RQ1 and RQ2 were investigated through textual analyses using methodology of corpus linguistics, whereby a corpus-based approach was adopted utilising Gardner and Davies's (2014) New Academic Vocabulary List (AVL) as the basis for identification of academic vocabulary in the students' writing. The identified academic vocabulary items were subsequently explored from the perspective of density (i.e. coverage) and diversity (i.e. variety). The textual data were complemented by interview and survey data collected to address RQ3.

The findings generated in response to RQ1 showed the impact of academic vocabulary items contained in the assignment brief on both the density as well as diversity of academic vocabulary across the four writing genres of assessed academic writing investigated in this study. The impact of the assignment brief was found to be particularly noticeable in students less confident in using appropriate vocabulary, their own words or paraphrasing, who tended to draw on the assignment brief as a vocabulary repository when completing their written assignments. A further factor impacting both the density and diversity of academic vocabulary was the writing genre; however, the genre's impact was found to be less significant than that of the assignment brief and topic. The assignment topic was shown to be a factor impacting primarily the density of academic vocabulary with academic vocabulary items contained in the assignment titles forming a considerable proportion of all academic vocabulary in the students' texts. This finding also showed the different behaviour of topic-specific words in different contexts and disciplines, underlining the importance of knowledge of the meaning and collocational preferences of individual words in

specific contexts. Another finding related to a small core of productive academic vocabulary types shared by the majority of students. These core academic vocabulary items were found to display different functions across the writing genres, which was attributed to the different social purposes of individual genres.

In terms of RQ2 investigating the development of academic vocabulary over an academic year, measured on the basis of the first and last written assignment (i.e. Exposition and Discussion Essay), the findings showed an overall increase in both the density and diversity of academic vocabulary. This finding thus shows that the period of an academic year (i.e. 24 teaching weeks delivered over six months in the present study) was sufficient for novice student writers at a foundation-level of study to increase their productive academic vocabulary from the perspectives of density and diversity. This increase was, however, not distributed across all topics equally as some topics were found to have prompted a denser and more diverse usage of academic vocabulary than others. Further, the increase in the density of academic vocabulary was shown to be impacted by increased density of academic vocabulary items drawn from the assignment brief, suggesting that students may have been integrating these academic vocabulary items in their assignments to a greater extent as they progressed so as to explicitly address the assignment instructions. In contrast, the increase in the diversity of academic vocabulary was found to be related to drawing on a larger number of reading sources. Further changes were also identified in the functions of the core academic vocabulary, which was attributed to the genres' social purposes. In addition, the students were found to deploy different metadiscourse markers and the use of nominal style also became more prevalent in the final written assignment.

RQ3 was concerned with the students' perceptions of the factors impacting their academic vocabulary learning and was addressed by interview and survey data. These two types of data highlighted several factors that the students perceived as impacting their acquisition, deployment and development of academic vocabulary. Among these were the instructed environment and exposure to appropriate reading and reference sources, which were found to play a crucial role in the acquisition of academic vocabulary. Textual sources such as teaching materials, assignment briefs and reading texts, were reported to play a role in the deployment of academic vocabulary in the students' assessed academic writing. As far as the development of

academic vocabulary was concerned, the main contributing factors reported by the students included opportunities for practice together with receiving feedback on written production and peers.

8.2 Contributions of study

This study's contribution lies in adding to the existing body of research on student writing (8.2.1) as well as to methodology (8.2.2). The findings generated by the present study also have potentially important implications for pedagogy relating to IFP contexts (8.2.3).

8.2.1 Contributions to research

The current study has contributed to the body of research that exists in the domain of university student writing, particularly with regard to the usage of academic vocabulary which is considered a key aspect of writing in EAP contexts. Despite an extensive body of literature on the deployment of academic vocabulary in learner writing conducted in diverse educational contexts and at different levels of study (e.g., Brun-Mercer & Zimmerman, 2015; Cons, 2012; Coxhead, 2012; Csomay & Prades, 2018; Durrant, 2016; Knoch et al., 2015, 2014; Masrai & Milton, 2018; Nadarajan, 2011; Olinghouse & Wilson, 2013; Storch, 2009; Storch & Tapper, 2009; Xudong et al., 2010) (3.2.5), the review of the literature has shown that the academic vocabulary in the writing produced by foundation-level students remains an under-researched area (3.2.6). This study has, therefore, addressed this important omission in the current body of research by providing insights into various aspects of the usage of academic vocabulary in academic writing produced by a multilingual group of students at a foundation-level of study in the context of a UK university, thereby making several contributions to vocabulary and writing research.

Specifically, this work has extended our knowledge of the deployment of academic vocabulary across four genres of assessed academic writing (i.e. Exposition Essays, Problem Questions, Research Reports and Discussion Essays) produced by international foundation students in the context of a generic IFP. This can be seen as an important contribution to the current body of literature considering the vital role that academic vocabulary knowledge and genre awareness play in written production in university contexts (1.2). In addition, there is a lack of studies exploring the deployment of academic vocabulary across different writing genres and in assessed

academic writing. With the exception of a few studies (Csomay & Prades, 2018; Durrant, 2016), most studies into academic vocabulary in university student writing do not consider different writing genres and use unassessed academic writing. This omission in the current body of research has thus been addressed by the present study.

Further, this study contributes to our understanding of the changes in the deployment of academic vocabulary over the duration of an academic year by offering insights into the development of productive academic vocabulary from a longitudinal perspective. Hence, this study also contributes to academic vocabulary research conducted over a period of time, which lacks studies in the context of foundation-level provision since all identified longitudinal studies into academic vocabulary were carried out at higher levels of degree study. In addition, none of the identified studies were set in the context of UK tertiary education as the majority of them were conducted at Australian universities (Knoch et al., 2015, 2014; Storch, 2009; Storch & Tapper, 2009). Further, with the exception of Knoch et al.'s (2015, 2014) studies, these studies measured the development of academic vocabulary over shorter periods of time (e.g. 10 - 12 weeks).

Next, the findings in this study provide an understanding of the acquisition, deployment and development of academic vocabulary from the learners' perspective, which is also under-reported in the current body of research into academic vocabulary. Although some studies into academic vocabulary (Coxhead, 2012) consider students' perceptions, these related to different factors from those explored in the present study.

The present study, therefore, contributes in several ways to our understanding of the assessed academic writing of foundation-level students from different cultural, educational and linguistic backgrounds characteristic of UK universities, from the perspective of academic vocabulary as the building blocks of academic texts. As well as adding to the existing body of research, the present study has also made several methodological contributions, outlined next.

8.2.2 Methodological contributions

Further contribution relates to the methodology employed in the current study, specifically the use of a learner corpus comprising over a million running words of

multilingual international foundation-level students' assessed academic writing; that is, a corpus considerably larger and in terms of the students' L1s (representing 55 linguistic backgrounds) more heterogeneous than most learner corpora compiled for the investigation of student writing. The composition of the learner corpus in this study is believed to have provided a suitable sample for the investigation of intergroup homogeneity in foundation-level students' assessed writing (Crossley & McNamara, 2011) (4.3.3). This approach can be regarded as important since it enabled examination of the deployment of academic vocabulary of international foundation students in the context of a UK university characterised by a diverse student population (1.1).

Next, an academic word list has not previously been utilised to investigate academic vocabulary in the different genres of assessed academic writing produced by foundation-level students in the context of a UK university. Further, the triangulation of data, whereby the corpus-based analysis of textual data was complemented by interview and survey data, can also be seen as a methodological contribution as it offered insights not only into the writing product typically focused on in the writing as completed activity approach to writing (Hyland, 2016b) (4.2.1) enabled by corpus linguistics (2.2.3), but also into the writing processes employed by the participants, prevalent in the academic literacies approach to investigation of student writing (2.3).

A further methodological contribution can be seen in the number of texts analysed in this study, which was found to correlate with internal variation, which in turn seemed to impact the density and diversity of academic vocabulary. Next, the investigation of variation between writing genres highlighted the importance of attention being given to the variation within writing genres as well, which may be impacted by factors other than the writing genre itself, such as the topic or discipline.

In addition to the above contributions to methodology, the present study has yielded some potentially important implications for pedagogy, considered next.

8.2.3 Implications for pedagogy

The findings reported in this study highlight several factors that played an important role in the acquisition, deployment and development of academic vocabulary in the context of multilingual novice student writers in the context of a generic IFP. Given

that written assignments are the main form of assessment at a tertiary level of education (1.2) and considering the importance of academic vocabulary in academic writing (3.2.3), these findings have potentially important implications for classroom practices in relation to IFP provision catering for a socially, ethnically and linguistically diverse student population, discussed next.

The role of the assignment brief

Drawing on the vocabulary items contained in the assignment brief was found to be a compensatory strategy of less proficient or confident novice writers as well as a way of explicitly addressing the assignment instructions. Practitioners thus ought to be aware of the important role that the assignment brief plays in student written production in terms of vocabulary usage, where it serves not only as a pedagogical tool (Russell, 2001) in a form of vocabulary repository (Plakans & Gebril, 2012) providing students with appropriate vocabulary to integrate in their writing, but also as a strategy of responding to the assignment instructions. Accordingly, practitioners should include such vocabulary in the assignment brief that the students would be expected to deploy in their written assignments. Learners' attention should also be drawn to the important role that the wording of the assignment brief plays not only in addressing the assignment instructions, but they should also be made aware that the brief can be used as one of the textual input sources that they can draw on during the process of completing their written assignments.

The role of the topic

The impact of the assignment topic highlighted the importance of knowledge of topic-specific vocabulary (Olinghouse & Wilson, 2013) and the meaning and collocational preferences of individual words in specific contexts (Hyland & Tse, 2007, 2009). Some topics were also found to prompt a denser and more diverse usage of academic vocabulary in student written production. The pedagogical implication of this finding lies in explicit focus on topic-specific vocabulary items and the various word combinations in which they tend to co-occur. However, practitioners should also be conscious of the fact that each text is likely to contain a high number of topic-specific vocabulary, which may be less useful for learners in more general academic contexts. Thus, it is important to find a balance between helping learners acquire topic-related vocabulary necessary for a specific task and more general academic vocabulary commonly used in various academic contexts. Further, practitioners

ought to be aware that not all topics may give students an equal opportunity for academic vocabulary integration (Knoch et al., 2015, 2014).

The role of the writing genre

The writing genre and its social purpose was found to impact the function and meaning of academic vocabulary items in their co-textual environments, including the changes in the functions of the core academic vocabulary. Practitioners, therefore, ought to explicitly focus on the characteristics of the different writing genres that the students are required to produce (Nesi & Gardner, 2012). In particular, students' attention ought to be drawn to the genres' social purpose and how this in turn impacts vocabulary selection to fulfil this purpose (Csomay & Prades, 2018; Durrant, 2016; Olinghouse & Wilson, 2013); in other words, learners should be made aware of the type of academic vocabulary appropriate to deploy in different writing genres. One of the important textual features arising from the genre's social purposes relates to metadiscourse. This aspect of academic writing should, therefore, not be overlooked as effective teaching of metadiscourse helps students develop an awareness of audience and equips them with the means to appropriately engage with that audience (Hyland, 2005, p. 181).

The role of reading and reference sources

Drawing on a wider range of reading sources was found to lead to an increased diversity of academic vocabulary. Reading sources were also reported as one of the main contributing factors impacting the acquisition of academic vocabulary. Practitioners thus ought to encourage and facilitate exposure to suitable academic texts and emphasise the crucial role that reading plays in the writing process where reading materials serve as a source of both information and appropriate vocabulary (Plakans & Gebril, 2012). In particular, learners should be provided with relevant reading materials that would enable them to notice new academic vocabulary (Nation, 2001) and should also have opportunities to work with academic texts in a classroom under teachers' guidance aimed at equipping learners with relevant learning strategies, such as identifying unfamiliar academic vocabulary and how these newly-met vocabulary items are used in meaningful contexts. This would enable them to select relevant vocabulary items for integration in their written production (Coxhead, 2012; Gebril & Plakans, 2016) since following a text as a model can help learners feel more confident in deploying academic vocabulary in their

writing (Cons, 2012). The use of reading sources was related to the reported reliance on quotations, considered a helpful strategy particularly for novice writers (Coxhead, 2012). Hence, practitioners could also focus on the use of quotations as a way of incorporating appropriate vocabulary in the students' written assignments in academically accepted ways without over-using or over-relying on them, perhaps in conjunction with paraphrasing techniques.

Various reference materials were also reported as a source of new academic vocabulary, which can benefit learners by providing access to a large number of words and their meanings together with a great deal of information about them often with examples of usage in their contextual environments including multi-word units such as collocations, which are regarded as important in vocabulary learning in general (Nation, 2001) and in academic contexts in particular (Hyland, 2008; Martinez & Schmitt, 2012; Simpson-Vlach & Ellis, 2010; Ackermann & Chen, 2013; Staples et al., 2013). Based on the students' perceived benefits of drawing on reference sources, a certain amount of classroom time should thus be dedicated to introducing learners to suitable reference materials as well as to assisting them in using these reference sources effectively and independently in both reading comprehension and written production (Salehi & Habibi, 2015; Schmitt, 1997).

The role of instructed environment and teaching materials

Instructed environment (i.e. teaching, lectures, seminars) was reported as a perceived contributing factor impacting the acquisition of academic vocabulary, highlighting the crucial role that practitioners and instruction play in the students' learning of academic vocabulary (Cons, 2012; Corson, 1997; Nation, 2001; Storch & Tapper, 2009). Based on the students' perceptions, practitioners should ensure that sufficient time is dedicated to explicit focus on academic vocabulary items in a classroom. This should involve not only introducing new academic vocabulary, but also "explicit instruction on how to effectively use academic words in writing" (Cons, 2012, p. 630), whereby modelling in relation to academic vocabulary deployment is an integral part of an effective lesson (Coxhead, 2012), particularly in relation to the various writing genres that students are required to produce (Olinghouse & Wilson, 2013). This is particularly important given the impact of the writing genre on the function, meaning and collocational behaviour of academic vocabulary (Hyland & Tse, 2007, 2009). Further, considering the important role of the nominal style in academic

writing (2.2.3), there should be an explicit focus on the importance of nouns in the academic register. This should involve sufficient practice with a particular focus on the specific structures functioning as noun modifiers and demonstrations of the process of word transformations using the suffixes characteristic of nominalisation and the subsequent changes in the sentence structure resulting from other word classes being turned into noun phrases (2.2.3).

The role of instructed environment is closely related to the role of teaching materials, reported by the participants as an important contributor to the learning of academic vocabulary. Practitioners should, therefore, pay particular attention to the content of the teaching materials which they create and present learners with. Specifically, teaching materials should provide learners with the conditions necessary for vocabulary learning, which include noticing, retrieval and generation (3.2.2). That is, teaching materials should introduce learners to new academic vocabulary and serve as model language providing specific examples of usage in various authentic contexts (the noticing stage). Further, teaching materials should enable learners to practise newly met academic vocabulary (the retrieval stage), followed by opportunities to use the acquired academic vocabulary in production in a form of meaning-focused output activities (the generation stage).

The role of peer support

Peer support was noted by the participating students as a contributor to the improvement of academic vocabulary. Thus, given the reported value of learning from peers, practitioners should not only make learners aware of the benefits of peer support and group work in the process of academic vocabulary learning, but should also facilitate group working opportunities for learners during as well as outside lessons since it is considered a valuable social discovery and consolidation vocabulary learning strategy (Schmitt, 1997) (3.2.2). Some of the activities promoting the acquisition, deployment and development of productive academic vocabulary, in which the learners could be involved, include collaborative writing tasks (Jelodar & Farvardin, 2019; Swain, 2001), for example. This type of activity can be seen as beneficial for the learners' academic vocabulary development as it covers all three stages of vocabulary learning (i.e. noticing, retrieval and generation) and relates to social discovery and consolidation vocabulary learning strategies. During these activities, the group composition relating to the learners' linguistic background and

linguistic competence should also be considered to maximise the benefit of group work. That is, learners with different L1s ought to be encouraged to work together to promote the use of English only, and students with different levels of linguistic proficiency should be grouped together to give less proficient students the opportunity to learn from their more capable peers, in line with the notion of 'zone of proximal development' (Vygotsky, 1978).

The role of practice and feedback

Being provided with opportunities for practice was perceived by the participants as one of the factors contributing to their development of academic vocabulary. The learners' perspectives thus highlight the importance of giving students ample opportunities to practise the integration of academic vocabulary in written production (Brun-Mercer & Zimmerman, 2015; Carlo et al., 2004; Cons, 2012; Coxhead, 2012; Nagy & Townsend, 2012), particularly in the specific genes that they are required to produce (Olinghouse & Wilson, 2013). Writing should thus be assigned frequently both during lessons under the teachers' guidance as well as independently by assigning writing tasks for learners to complete outside the classroom. In addition, given the perceived value of feedback, writing practice should always be accompanied by feedback on their written production in general and the usage of academic vocabulary deployment in particular, providing students with clear guidance on areas for further improvement as well as suggestions on how they can improve the language areas in question (Knoch, et al. 2015, 2014; Weaver, 2006). Considering that written assignments are a high-stakes activity in most university settings, practitioners should also encourage and enable learners to seek formative feedback on their assignment drafts prior to submission (Neumann, 2014; Wingate, 2010).

8.3 Limitations of study

Despite the potentially important contributions of this study to EAP vocabulary research and pedagogy, several limitations need to be acknowledged. First, notwithstanding a relatively large number of participants across three campuses in different parts of the world, the sample is limited to one university and is based in the context of a generic (as opposed to discipline-specific) IFP. Hence, the findings may not be generalisable to other IFPs which may differ in the delivery of the programme, IFPs preparing students for a specific discipline of study, or wider EAP

contexts. Further limitations could be seen in the methodology employing a corpus-based approach. One of the potential weaknesses of the corpus-based approach adopted in this study may lie in drawing on a list of pre-existing academic vocabulary, which might have resulted in an omission of some other vocabulary items that the students may have acquired and deployed in their written production. Another limitation relating to the corpus-based methodology can be seen in the fact that the focus of this study was primarily on the writing product, neglecting the dynamic interaction between the students and the social contexts within which knowledge is constructed. To address this short-coming and to complement the textual analysis, interviews and surveys were used to gain insights into the learners' experiences. However, the limitations of these methods also need to be considered. As far as interviews are concerned, one of the weaknesses relates to the issue of reactivity (4.3.5) and the students' lack of awareness of the various factors that may have played a contributory role in the acquisition, deployment and development of their academic vocabulary. In addition, due to the timing of the data collection, it was not possible to interview the participants after submission of each of the collected assignments, which might have offered greater insights into the strategies employed by the students in terms of the acquisition and subsequent integration of academic vocabulary in their written production. As far as the questionnaires are concerned, their potential weakness lies in the limited knowledge on the researcher's part as to the genuineness of the participants' responses. Further, due to the relatively large sample (N = 193) generating a large number of academic vocabulary items, it was not possible to explore the identified most frequently-used academic vocabulary items in their contextual environment for appropriacy of usage, which could potentially have provided insights into the students' overall writing quality and improvement. It was also beyond the scope of this study to investigate other factors that may have impacted the deployment and development of the students' academic vocabulary in their written production, such as student motivation, level of integration, cognitive processes involved in writing tasks, teaching practices or language and literacies outside the university settings, which could provide further valuable insights into the factors that impact academic vocabulary learning and production.

8.4 Future research

Future research would benefit from studies into the deployment and development of academic vocabulary in foundation-level students' assessed writing at a greater number of universities and on discipline-specific IFPs. Considerably more research into productive academic vocabulary is also needed in other EAP contexts to determine whether this study's findings are generalisable across various EAP settings. Accompanying textual analysis of every collected piece of student writing with interview data, particularly in longitudinal studies, could also be usefully explored. This would offer valuable insights into the various writing strategies used by the students, including textual borrowing, for instance. This is an area of student writing research which requires further investigation as "research on source-based writing and lexical borrowing is still in its infancy", meaning that several questions still remain unanswered with regard to how writers integrate vocabulary borrowed from source texts and whether the topics they are writing on impact the lexical quality of their writing (Gebril & Plakans, 2016, p. 87). In terms of source texts and textual borrowing, further studies regarding the role that the assignment brief plays in the process of completing written assignments would be worthwhile together with further research into the role of the assignment topic. In addition, the relationship between the usage of academic vocabulary in written assignments and student writing quality and improvement could also be investigated.

To further complement textual analysis of academic vocabulary displayed in students' texts and to obtain a more accurate assessment of the development of the students' academic writing skills and the strategies they employ, it might be useful to focus on other factors potentially impacting students' productive knowledge of academic vocabulary such as their level of integration, motivation, cognitive processes, teaching practices or literacies outside university contexts. Useful insights could also be gained from exploring the influence of students' prior experiences and educational backgrounds on their writing practices, including the ways in which novice students' writing practices are impacted by university entry exams such as IELTS.

References

Ackermann, K., & Chen, Y. H. (2013). Developing the Academic Collocation List (ACL) - A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes*, *12*(4), 235–247.

Adams, J. (2015). Student-writing tutors: Making sense of 'academic literacies'. In T. Lillis, K. Harrington, M. R. Lea, & S. Mitchell (Eds.), *Working with academic literacies: Case studies towards transformative practice* (pp. 65–73). Parlor Press.

Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, *31*(2), 81–92.

Alsop, S., & Nesi, H. (2009). Issues in the development of the British Academic Written English (BAWE) corpus. *Corpora*, *4*(1), 71–83.

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, *57*(5), 13–18.

Anthony, L. (2013). AntWordProfiler. Tokyo, Japan: Waseda University. Retrieved from http://www.laurenceanthony.net/software

Anthony, L. (2015). TagAnt. Tokyo, Japan: Waseda University. Retrieved from http://www.laurenceanthony.net/software

Anthony, L. (2018). AntConc. Tokyo, Japan: Waseda University. Retrieved from http://www.laurenceanthony.net/software

Baker, M. (1988). Sub-technical vocabulary and the ESP teacher: an analysis of some rhetorical items in medical journal articles. *Reading in a Foreign Language*, *4*(2), 91–105.

Baker, P., & Egbert, J. (Eds.). (2016). *Triangulating methodological approaches in corpus-linguistic research*. Routledge.

Banks, D. (2005). On the historical origins of nominalized process in scientific text. *English for Specific Purposes*, *24*(3), 347–357.

Baratta, A. M. (2010). Nominalization development across an undergraduate academic degree program. *Journal of Pragmatics*, *42*(4), 1017–1036.

Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography*, *6*(4), 253–279.

Baumann, J. F., & Graves, M. F. (2010). What is academic vocabulary? *Journal of Adolescent & Adult Literacy*, 4–12.

Bawarshi, A. S., & Reiff, M. J. (2010). *Genre: An introduction to history, theory, research, and pedagogy*. Parlor Press.

Bhatia, V. K. (1995). Genre analysis and ESP. *The Journal of TESOL France*, *2*(2), 161–179.

Bhatia, V. K. (1997). Introduction: Genre analysis and world Englishes. *World Englishes*, *16*(3), 313–319.

Bhatia, V. K. (2004). *Worlds of written discourse: A genre-based view*. Continuum.

Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.

Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, *8*(4), 243–257.

Biber, D. (2019). Text-linguistic approaches to register variation. *Register Studies*, *1*(1), 42–75.

Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge University Press.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371–405.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, *45*(1), 5–35.

Biber, D., Gray, B., & Poonpon, K. (2013). Pay attention to the phrasal structures: Going beyond t-units — a response to WeiWei Yang. *TESOL Quarterly*, *47*(1), 192–201.

Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*, *37*(5), 639–668.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.

Blagojević, S. (2016). Book review. *Journal of English for Academic Purposes*, *24*, 75–77.

Bourdieu, P., Passeron, J.-C., & Saint Martin, M. (1994). *Academic discourse*. Stanford University Press.

Bourke, J. M. (2005). The grammar we teach. *Reflections on English Language Teaching*, *4*(2), 85–97.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge University Press.

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the New General Service List. *Applied Linguistics*, *36*(1), 1–22.

British Association for Applied Linguistics (BAAL). (2016). *Recommendations on good practice in applied linguistics*. Retrieved from https://baalweb.files.wordpress.com/2016/10/goodpractice_full_2016.pdf

British Education Research Association (BERA). (2018). *Ethical guidelines for educational research*. Retrieved from https://www.bera.ac.uk/researchers-resources/publications/ethical-guidelines-for-educational-research-2018

Browne, C. (2013). The New General Service List: Celebrating 60 years of vocabulary learning. *The Language Teacher*, *37*(4), 13–16.

Browne, C., Culligan, B., & Phillips, J. (2013a). New Academic Word List. Retrieved from http://www.newgeneralservicelist.org/nawl-new-academic-word-list

Browne, C., Culligan, B., & Phillips, J. (2013b). New General Service List. Retrieved from http://www.newgeneralservicelist.org/

Brun-Mercer, N., & Zimmerman, C. B. (2015). Fostering academic vocabulary use in writing. *The CATESOL Journal*, *27*(1), 131–148.

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, *26*, 42–65.

Bunton, D. (2002). Generic moves in PhD thesis introductions. In J. Flowerdew (Ed.), *Academic discourse* (pp. 57–75). Longman.

Bunton, D. (2005). The structure of PhD conclusion chapters. *Journal of English for Academic Purposes*, *4*, 207–224.

Cambridge Dictionary. (2020). Retrieved from https://dictionary.cambridge.org/

Camp, H. (2012). The psychology of writing development - and its implications for assessment. *Assessing Writing*, *17*(2), 92–105.

Carlo, M. S., August, D., Mclaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., Lively, T. J., & White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, *39*(2), 188–215.

Carter, R. (1992). *Vocabulary: Applied linguistic perspectives*. Routledge.

Carver, R. P. (1994). Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: Implications for instruction. *Journal of Reading Behavior*, *26*(4), 413–437.

Chang, A. C.-S., & Hu, H. M. (2018). Learning vocabulary through extensive reading: Word frequency levels and L2 learners' vocabulary knowledge level. *Teaching English as a Second or Foreign Language*, *22*(1), 1–21.

Chang, C.-S. A. (2012). Gains to L2 learners from extensive listening: listening development, vocabulary acquisition and perceptions of the intervention. *Hong Kong Journal of Applied Linguistics*, *14*(1), 25–47.

Chang, G. C. (2014). Writing feedback as as exclusionary practice in higher educaiton. *Australian Review of Applied Linguistics*, *37*(3), 262–275.

Chanock, K., Whitmore, S., & Nishitani, M. (2015). Disciplined voices, disciplined feelings: Exploring constraints and choices in a thesis writing circle. In T. M. Lillis, K. Harrington, M. R. Lea, & S. Mitchell (Eds.), *Working with academic literacies: Case studies towards transformative practice* (pp. 106–115). Parlor Press.

Chen, C. N., Chen, S. C., Chen, S. H. E., & Wey, S. C. (2013). The effects of extensive reading via e-books on tertiary level EFL students' reading attitude, reading comprehension and vocabulary. *The Turkish Online Journal of Educational Technology*, *12*(2), 303–312.

Christie, H. (2007). Higher education and spatial (im)mobility: Nontraditional students and living at home. *Environment and Planning A*, *39*(10), 2445–2463.

Chuang, F.-Y., & Nesi, H. (2006). An analysis of formal errors in a corpus of Chinese student writing. *Corpora*, *1*(2), 251–271.

Coffin, C., & Donohue, J. P. (2012). Academic literacies and systemic functional linguistics: How do they relate? *Journal of English for Academic Purposes*, *11*(1), 64–75.

Cohen, A., Glasman, H., Rosenbaum-Cohen, P. R., Ferrara, J., & Fine, J. (1979). Reading English for Specialized Purposes: Discourse analysis and the use of student informants. *TESOL Quarterly*, *13*(4), 551–564.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cons, A. M. (2012). The use and misuse of academic words in writing: Analyzing the writing of secondary English learners and redesignated learners. *TESOL Journal*, *3*(4), 610–638.

Corson, D. (1997). The learning and use of academic English words. *Language Learning*. 47(4), 671–718.

Cowan, J. R. (1974). Lexical and syntactic research for the design of EFL reading materials. *TESOL Quarterly*, *8*(4), 389–399.

Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, *34*(2), 213–238.

Coxhead, A. (2012). Academic vocabulary, writing and English for Academic Purposes: Perspectives from second language learners. *RELC Journal*, *43*(1), 137–145.

Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing*, *16*(3), 129–147.

Coxhead, A., & Hirsch, D. (2007). A Pilot Science-Specific Word List. *Revue Française de Linguistique Appliquée*, *XII*, 65–78.

Coxhead, A., & Nation, P. (2001). The specialised vocabulary of English for academic purposes. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for Academic Purposes* (pp. 252–267). Cambridge University Press.

Crossley, S. A., & McNamara, D. S. (2011). Shared features of L2 writing: Intergroup homogeneity and text classification. *Journal of Second Language Writing*, *20*(4), 271–285.

Csomay, E., & Prades, A. (2018). Academic vocabulary in ESL student papers: A corpus-based study. *Journal of English for Academic Purposes*, *33*, 100–118.

Dang, T. N. Y. (2018a). A Hard Science Spoken Word List. *International Journal of Applied Linguistics*, *169*(1), 44–71.

Dang, T. N. Y. (2018b). The nature of vocabulary in academic speech of hard and soft-sciences. *English for Specific Purposes*, *51*, 69–83.

Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The Academic Spoken Word List. *Language Learning*, *67*(4), 959–997.

de Oliveira, L. C. (2015). A systemic-functional analysis of english language learners' writing. *Documentacao de Estudos Em Linguistica Teorica E Aplicada (DELTA)*, *31*(1), 207–237.

de Vries, D. F., & Raffin, V. (2016). Specialist knowledge: An interactive approach. *InForm*, (16), 11–13.

Devitt, A. J. (2015). Genre performances: John Swales' genre analysis and rhetorical-linguistic genre studies. *Journal of English for Academic Purposes*, *19, 44-51*.

Díaz-Negrillo, A., & Thompson, P. (2013). Learner corpora: Looking towards the future. In A. Diaz-Negrillo, N. Ballier, & P. Thompson (Eds.), *Automatic Treatment and Analysis of Learner Corpus Data* (pp. 9–30). John Benjamins.

Dobao, A. F. (2014). Vocabulary learning in collaborative tasks: A comparison of pair and small group work. *Language Teaching*, *18*(4), 497–520.

Donley, K. M., & Reppen, R. (2001). Using corpus tools to highlight academic vocabulary in SCLT. *TESOL Journal*, *10*(2/3), 7–12.

Dougherty, E. (2012). *Assignments matter: Making the connections that help students meet standards*. Association for Supervision Curriculum Development.

Dudley-Evans, T. (1995). Genre models for the teaching of academic writing to second language speakers: Advantages and disadvantages. *The Journal of TESOL France*, *2*(2), 181–192.

Durrant, P. (2014). Discipline and level specificity in university students' written vocabulary. *Applied Linguistics*, *35*(3), 328–356.

Durrant, P. (2016). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes*, *43*, 49–61.

Durrant, P., & Mathews-Aydinli, J. (2011). A function-first approach to identifying formulaic language in academic writing. *English for Specific Purposes*, *30*(1), 58–72.

Egbert, J., & Gray, B. (2019). Editorial: Register and register variation. *Register Studies*, *1*(1), 1–9.

Eggins, S. (2004). *An introduction to systemic functional linguistics* (2nd ed.). Continuum International Publishing Group.

Ellis, R. (1999). Factors in the incidental acquisition of second language vocabulary from oral input. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 35–61). John Benjamins Publishing Company.

Engels, L. K. (1968). The fallacy of word-counts. *International Review of Applied Linguistics in Language Teaching*, *6*(3), 213–231.

Farrell, P. (1990). *Vocabulary in ESP: A lexical analysis of the English of electronics and a study of semi-technical vocabulary*. *CLCS Occasional Paper No. 25*. Dublin.

Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Lawrence Erlbaum Associates.

Fisher, D., & Frey, N. (2008). *Word wise and content rich: Five essential steps to teaching academic vocabulary*. Heinemann.

Flowerdew, J. (2002). Genre in the classroom: A linguistic approach. In A. M. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 91–102). Lawrence Erlbaum Associates.

Flowerdew, J. (2015). John Swales's approach to pedagogy in genre analysis: A perspective from 25 years on. *Journal of English for Academic Purposes*, *19*, 102–112.

Flowerdew, J. (2016). English for Specific Academic Purposes (ESAP) writing: Making the case. *Writing & Pedagogy*, *8*(1), 5–32.

Flowerdew, J., & Peacock, M. (2001). Issues in EAP: A preliminary perspective. In J. Flowerdew & M. Peacock (Eds.), *Research Perspectives on English for Academic Purposes* (pp. 8–24). Cambridge University Press.

Flowerdew, L. (2003). A combined corpus and systemic-functional analysis of the problem-solution pattern in a student and professional corpus of technical writing. *TESOL Quarterly*, *37*(3), 489–511.

Gardner, D., & Davies, M. (2014). A New Academic Vocabulary List. *Applied Linguistics*, *35*(3), 305–327.

Gardner, S., & Nesi, H. (2013). A classification of genre families in university student writing. *Applied Linguistics*, *34*(1), 25–52.

Gardner, S., Nesi, H., & Biber, D. (2018). Discipline, level, genre: Integrating situational perspectives in a new MD analysis of university student writing. *Applied Linguistics*, 1–30.

Gardner, S., & Powell, L. (2006). *An investigation of genres of assessed writing in British higher education: a Warwick-Reading-Oxford Brookes project* (Research, scholarship and practice in the area of Academic Literacies Annual one-day seminar).

Gebhard, M. (2010). Teacher education in changing times: A systemic functional linguistics (SFL) perspective. *TESOL Quarterly*, *44*(4), 797–803.

Gebril, A., & Plakans, L. (2016). Source-based tasks in academic writing assessment: Lexical diversity, textual borrowing and proficiency. *Journal of English for Academic Purposes*, *24*, 78–88.

Gray, B. (2016). Lexical Bundles. In P. Baker & J. Egbert (Eds.), *Triangulating methodological approaches in corpus-linguistic research* (pp. 33–56). Routledge.

Groves, M. (2016). Providing focus, or limiting options? Lexis, disciplines and English for Academic Purposes. *InForm*, (16), 25–26.

Gurr, H. (2016). How I survived teaching English for Mathematics. *InForm*, (16), 23–24.

Gustafson-Pearce, O. (2009). The assignment brief. *Procedia Social and Behavioral Sciences*, (World Conference on Educational Sciences), 2584–2588.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. Edward Arnold.

Halliday, M. A. K. (2004). *The language of science*. Continuum.

Halliday, M. A. K. (2014). *Halliday's introduction to functional grammar*. (C. M. I. M. Matthiessen, Ed.) (4th ed.). Routledge.

Halliday, M. A. K., & Martin, J. R. (1993). *Writing science: Literacy and discursive power*. University of Pittsburgh Press.

Hammersley, M. (2009). *MRes programme Module A: An introduction to social research.* The Open University.

Hammersley, M., & Atkinson, P. (2007). *Ethnography: Principles in practice* (3rd ed.). Routledge.

Harmon, J. M., Wood, K. D., & Medina, A. L. (2009). Vocabulary learning in the content areas: Research-based practices for middle and secondary school classrooms. In K. D. Wood & W. E. Blanton (Eds.), *Literacy instruction for adolescents: Research-based practice* (pp. 344–367). The Guilford Press.

Heaviside, H. (2017). Detangling the web of methodology: A doctoral student's perspective on chaos and contradiction. *Sport & Exercise Psychology Review*, *13*(1), 73–80.

HESA. (n.d.). Higher Education Student Data. Retrieved from https://www.hesa.ac.uk/data-and-analysis/students

Hiebert, E. H., & Lubliner, S. (2008). The nature, learning, and instruction of general academic vocabulary. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about vocabulary instruction* (pp. 106–129). International Reading Association.

Hilsdon, J., Syska, A., Hagyard, A., Eccles, S., Howells-, C., Sedghi, G., Morley, C., Drew, C., Malone, C., Loughlin, E., Briggs, S., Kavanagh, M., Brewer, S., & Wrigglesworth, J. (2019). Editorial. *Journal of Learning Development in Higher Education*, (15), 1–7.

Hinkel, E. (2003). Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly*, *37*(2), 275–301.

Holton, M. (2018). Traditional or non-traditional students?: incorporating UK students' living arrangements into decisions about going to university. *Journal of Further and Higher Education*, *42*(4), 556–569.

Home Office. (2019). *Tier 4 of the points based system – Policy guidance*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/812143/T4_Migrant_Guidance_JUNE_2019_FINAL_v3.pdf

Hsu, W. (2011). A Business Word List for prospective EFL Business postgraduates. *The Asian ESP Journal*, *7*(4), 63–99.

Hsu, W. (2013). Bridging the vocabulary gap for EFL medical undergraduates: The establishment of a medical word list. *Language Teaching Research*, *17*(4), 454–484.

Hsu, W. (2014). Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes*, *33*, 54–65.

Huang, K. (2015). More does not mean better: Frequency and accuracy analysis of lexical bundles in Chinese EFL learners' essay writing. *System*, *53*, 13–23.

Hunston, S. (2013). Systemic functional linguistics, corpus linguistics, and the ideology of science. *Text and Talk*, *33*(4–5), 617–640.

Hunston, S., & Thompson, G. (Eds.). (2000). *Evaluation in text: authorial stance and the construction of discourse*. Oxford University Press.

Huong, H. L. P. (2006). Learning vocabulary in group work in vietnam. *RELC Journal*, *37*(1), 105–121.

Huong, L. P. H. (2018). A survey study on academic vocabulary learning strategies by EFL university students. *Journal of Language Teaching and Research*, *9*(5), 1009–1016.

Hutton, W. E. (2016). Using Moodle's glossary activity to promote the acquisition of historical and political terms on an International Foundation Programme. *InForm*, (16), 8–10.

Hyland, K. (1998). Persuasion and context : The pragmatics of academic metadiscourse. *Journal of Pragmatics*, *30*, 437–455.

Hyland, K. (2002). Genre: language, context and litteracy. *Annual Review of Applied Linguistics*, *22*, 113–135.

Hyland, K. (2003). *Second language writing*. Cambridge University Press.

Hyland, K. (2004). *Genre and second language writing*. The University of Michigan Press.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.

Hyland, K. (2008). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, *18*(1), 41–62.

Hyland, K. (2009). *Teaching and researching writing* (2nd ed.). Pearson Education Limited.

Hyland, K. (2013). Writing in the university: Education, knowledge and reputation. *Language Teaching*, *46*(1), 53–70.

Hyland, K. (2015). Genre, discipline and identity. *Journal of English for Academic Purposes*, *19*, 32–43.

Hyland, K. (2016a). General and specific EAP. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for Academic Purposes* (pp. 17–29). Routledge.

Hyland, K. (2016b). Methods and methodologies in second language writing research. *System*, 59, 116–125.

Hyland, K., & Hamp-Lyons, L. (2002). EAP: Issues and directions. *Journal of English for Academic Purposes*, *1*(1), 1–12.

Hyland, K., & Shaw, P. (2016). *The Routledge handbook of English for Academic Purposes*. Routledge.

Hyland, K., & Tse, P. (2007). Is there an 'Academic Vocabulary'? *TESOL Quarterly*, *41*(2), 235–253.

Hyland, K., & Tse, P. (2009). Academic lexis and disciplinary practice: Corpus evidence for specificity. *International Journal of English Studies*, *9*(2), 111–129.

Hyon, S. (1996). Genre in three traditions: Implications for ESL. *TESOL Quarterly*, *30*(4), 693–722.

Jelodar, Z. Y., & Farvardin, M. T. (2019). Effects of collaborative tasks on EFL learners' writing productions. *International Journal of Instruction*, *12*(1), 389–406.

Jordan, R. R. (2002). The growth of EAP in Britain. *Journal of English for Academic Purposes*, *1*(1), 69–78.

Kaufhold, K. (2015). Making sense of my thesis: Master's level thesis writing as constellation of joint activities. In T. Lillis, K. Harrington, M. R. Lea, & S. Mitchell (Eds.), *Working with academic literacies: Case studies towards transformative practice* (pp. 87–95). Parlor Press.

Khani, R., & Tazik, K. (2013). Towards the development of an academic word list for applied linguistics research articles. *RELC Journal*, *44*(2), 209–232.

Knapp, P., & Watkins, M. (2005). *Genre, text, grammar: technologies for teaching and assessing writing*. University of New South Wales Press.

Knoch, U., Rouhshad, A., Oon, S. P., & Storch, N. (2015). What happens to ESL students' writing after three years of study at an English medium university? *Journal of Second Language Writing*, *28*, 39–52.

Knoch, U., Rouhshad, A., & Storch, N. (2014). Does the writing of undergraduate ESL students develop after one year of study in an English-medium university? *Assessing Writing*, *21*, 1–17.

Konstantakis, N. (2007). Creating a Business Word List for teaching Business English. *Estudios de Linguistica Inglesa Aplicada (ELIA)*, *7*, 79–102.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press Inc.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the Input Hypothesis. *The Modern Language Journal*, *73*(4), 440–464.

Krashen, S. (2012). Direct instruction of academic vocabulary : What about real reading? *Reading Research Quarterly*, *47*(3), 233.

Kress, G. (2007). Thinking about meaning and learning in a world of instability and multiplicity. *Pedagogies. An International Journal*, *1*, 19–34.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Brown University Press.

Kvale, S. (2007). *Doing interviews*. SAGE.

Lantolf, J. P. (2000). *Sociocultural theory and second language learning*. Oxford University Press.

Laufer, B. (1997). What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 140–155). Cambridge University Press.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical density in FL written production. *Applied Linguistics*, *16*(3), 307–322.

Lea, M. R., & Street, B. V. (1998). Student sriting in higher education: An academic literacies approach. *Studies in Higher Education*, *23*(2), 157–172.

Lea, M. R., & Street, B. V. (2006). The 'Academic Literacies' model: Theory and applications. *Theory into Practice*, *45*(4), 368–377.

Leedham, M. (2006). ' Do I speak better ?' A longitudinal study of lexical chunking in the spoken language of two Japanese students. *The East Asian Learner*, *2*(2).

Leedham, M. (2012). Combining intuition with corpus linguistic analysis: A study of marked lexical chunks in four Chinese students' undergraduate assignments. *NJES Nordic Journal of English Studies*, *11*(3), 155–187.

Lei, L., & Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes*, *22*, 42–53.

Leki, I. (2006). 'You cannot ignore': L2 graduate students' response to discipline-based written feedback. In K. Hyland (Ed.), *Feedback in second language writing: Contexts and issues* (pp. 266–286). Cambridge University Press.

Leki, I., & Carson, J. (1997). 'Completely different worlds': EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly*, *31*(1), 39–69.

Lessard-Clouston, M. (2008). Strategies and success in technical vocabulary learning: students' approaches in one academic context. *Indian Journal of Applied Linguistics*, *34*(1–2), 31–63.

Lessard-Clouston, M. (2010). Theology lectures as lexical environments: A case study of technical vocabulary use. *Journal of English for Academic Purposes*, *9*(4), 308–321.

Li, J., & Schmitt, N. (2010). The development of collocation use in academic texts by advanced L2 learners: A multiple case study approach. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 23–46). Bloomsbury.

Lillis, T. (2003). Student writing as 'Academic Literacies': Drawing on Bakhtin to move from critique to design. *Language and Education*, *17*(3), 192–207.

Lillis, T. (2019). 'Academic iteracies': sustaining a critical space on writing in academia. *Journal of Learning Development in Higher Education*, (15), 1–18.

Lillis, T., & Scott, M. (2007). Defining academic literacies research: issues of epistemology, ideology and strategy. *Journal of Applied Linguistics and Professional Practice*, *4*(1), 5–32.

Lillis, T., & Tuck, J. (2016). Academic Literacies: a critical lens on writing and reading in the academy. In K. Hyland & P. Shaw (Eds.), *The Routledge Handbook of English for Academic Purposes. Routledge* (pp. 30–43). Routledge.

Lillis, T., & Turner, J. (2001). Student writing in higher education: Contemporary confusion, traditional concerns. *Teaching in Higher Education*, *6*(1), 57–68.

Lin, S. (2018). The effect of group work on English vocabulary learning. *Journal of Education and Learning*, *7*(4), 163–178.

Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes*, *39*, 1–11.

Liu, L., & Zhang, Y. (2012). A contrastive analysis of repetition strategies employed in Chinese EFL writing and Chinese-English translation. *International Communication Studies*, *21*(2), 111–130.

Manning, A. (2008). From the editorial board…. . *InForm*, (1), 1.

Martin, J. R. (1997). Analysing genre: Functional parameters. In F. Christie & J. R. Martin (Eds.), *Genre and institutions: Social processes in the workplace and school* (pp. 3–39). Continuum.

Martin, J. R. (2009). Genre and language learning: A social semiotic perspective. *Linguistics and Education*, *20*(1), 10–21.

Martin, J. R. (2016). Meaning matters: A short history of systemic functional linguistics. *Word*, *62*(1), 35–58.

Martínez, I. A., Beck, S. C., & Panza, C. B. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes*, *28*, 183–198.

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, *33*(3), 299–320.

Masrai, A., & Milton, J. (2018). Measuring the contribution of academic and general vocabulary knowledge to learners' academic achievement. *Journal of English for Academic Purposes*, *31*, 44–57.

McCarthy, M. (2019). *Why do we need EAP word lists and what can we do with them? BALEAP Conference: Innovation, exploration and transformation, University of Leeds 12th - 14th April 2019*. Leeds.

McCarthy, M., & Carter, R. (1997). Written and spoken vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (5th ed., pp. 20–39). Cambridge University Press.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. Routledge.

McLean, S. (2018). Evidence for the adoption of the flemma as an appropriate word counting unit. *Applied Linguistics*, *39*(6), 823–845.

Mei, W. S., & Allison, D. (2005). Evaluative expressions in analytical arguments: Aspects of appraisal in assigned English language essays. *Journal of Applied Linguistics*, *2*(1), 105–127.

Melka, F. (1997). Receptive vs. productive aspects of vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 84–102). Cambridge University Press.

Milton, J. (2001). *Elements of a written interlanguage: A computational and corpus-based study of institutional influences on the acquisition of English by Hong Kong Chinese students*. *Research reports* (Vol. 2). Hong Kong.

Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, *25*, 235–256.

Nadarajan, S. (2011). The challenges of getting L2 learners to use academic words in their writings. *Electronic Journal of Foreign Language Teaching*, *8*(2), 184–200.

Nagy, W. (1997). On the role of context in first- and second-laguage vocabulary learning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 64–83). Cambridge University Press.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, *47*(1), 91–108.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge University Press.

Ndoricimpa, C. (2019). Scaffolding ESL tertiary students' challenges with essay genre: A Systemic Functional Linguistics Perspective. *International Journal of Research in English Education*, *4*(4), 55–69.

Nesi, H. (2002). An English Spoken Academic Wordlist. In *Proceedings of the tenth EURALEX International Congress* (p. 351–357, Vol.1). Copenhagen, Denmark.

Nesi, H., & Gardner, S. (2012). *Genres across the disciplines: Student writing in higher education*. Cambridge University Press.

Nesi, H., & Gardner, S. (2018). The BAWE corpus and genre families classification of assessed student writing. *Assessing Writing*, *38*, 51–55.

Nesi, H., Gardner, S., Forsyth, R., Hindle, D., Wickens, P., Ebeling, S., Leedham, M., Thompson, P., & Heuboeck, A. (2005). Towards the compilation of a corpus of assessed student writing: An account of work in progress. In *Corpus linguistics conference series 1(1)* (pp. 1–18).

Nesi, H., Gardner, S., Thompson, P., Wickens, P., Forsyth, R., Heuboeck, A., Holmes, J., Hindle, D., Ebeling, S., Leedham, M., & Alsop, S. (2008). *An investigation of genres of assessed writing in British higher education: Full research report ESRC end of award report*. *RES-000-23-0800*. Swindon.

Neumann, H. (2014). Teacher assessment of grammatical ability in second language academic writing: A case study. *Journal of Second Language Writing*, *24*(1), 83–107.

O'Flynn, J. (2019). An Economics Academic Word List (EAWL): Using online resources to develop a subject-specific word list and associated teaching-learning materials. *Journal of Academic Language and Learning*, *13*(1), 28–87.

Olinghouse, N. G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing*, *26*, 45–65.

Oppenheimer, D., Zaromb, F., Pomerantz, J. R., Williams, J. C., & Park, Y. S. (2017). Improvement of writing skills during college: A multi-year cross-sectional and longitudinal study of undergraduate writing performance. *Assessing Writing*, *32*, 12–27.

Oxford Learner's Dictionaries. (2019). Retrieved from https://www.oxfordlearnersdictionaries.com/about/wordlists/opal

Paquot, M. (2010). *Academic vocabulary in learner writing: From extraction to analysis*. Continuum International Publishing Group.

Pearson, W. S. (2020). Mapping English language proficiency cut-off scores and pre-sessional EAP programmes in UK higher education. *Journal of English for Academic Purposes*, *45*, 1–11.

Plakans, L., & Gebril, A. (2012). A close investigation into source use in integrated second language writing tasks. *Assessing Writing*, *17*(1), 18–34.

Potter, S. (2006). *Doing postgraduaterResearch* (2nd ed.). SAGE Publication Ltd.

Prosser, M., & Webb, C. (1994). Relating the process of undergraduate essay writing to the finished product. *Studies in Higher Education*, *19*(2), 125–138.

Qin, J. (2014). Use of formulaic bundles by non-native English graduate writers and published authors in applied linguistics. *System*, *42*(1), 220–231.

Redpath, K. (2015). Collaborative vocabulary lesson design on the University of Edinburgh's International Foundation Programme. *InForm*, (15), 12–13.

Richards, J. C. (1974). Word lists: Problems and prospects. *RELC*, *2*, 69–84.

Rose, D., & Martin, J. R. (2012). *Learning to write, reading to learn: genre, knowledge and pedagogy in the Sydney school*. Equinox Publishing Ltd.

Russell, D. R. (2001). Where do the naturalistic studies of WAC/WID point? A research review. In S. McLeod, E. Miraglia, M. Soven, & C. Thaiss (Eds.), *WAC for the new millenium: Strategies for continuing writing across the curriculum programs* (pp. 259–298). National Council of Teachers of English.

Salehi, H., & Habibi, S. B. (2015). EFL learners' attitudes towards using thesaurus part of Microsoft Word in vocabulary acquisition. In *1st International Conference on Teaching & Learning (ICTL 2015)* (pp. 61–65). Langkawi, Malaysia.

Sapsford, R. (2007). *Survey research* (2nd ed.). SAGE.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 199–227). Cambridge University Press.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.

Schmitt, N., & McCarthy, M. (1997). Editor's comments - pedagogy section. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 321–326). Cambridge University Press.

Serrano, R., Tragant, E., & Llanes, Á. (2012). A longitudinal analysis of the effects of one year abroad. *The Canadian Modern Language Review*, *68*(2), 138–163.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3–4), 591–611.

Shrestha, P. (2020). *Dynamic assessment of students' academic writing: Vygotskian and systemic functional linguistic perspectives*. Springer.

Shrestha, P., & Coffin, C. (2012). Dynamic assessment, tutor mediation and academic writing development. *Assessing Writing*, *17*(1), 55–70.

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, *31*(4), 487–512.

Sökmen, A. J. (1997). Current trends in teaching second language vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 237–257). Cambridge University Press.

Soler-Monreal, C., Carbonell-Olivares, M., & Gil-Salom, L. (2011). A contrastive study of the rhetorical organisation of English and Spanish PhD thesis introductions. *English for Specific Purposes*, *30*(1), 4–17.

Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication*, *33*(2), 149–183.

Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for Academic Purposes*, *12*(3), 214–225.

Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing*, *32*, 17–35.

Storch, N. (2009). The impact of studying in a second language (L2) medium university on the development of L2 writing. *Journal of Second Language Writing*, *18*, 103–118.

Storch, N., & Tapper, J. (2009). The impact of an EAP course on postgraduate writing. *Journal of English for Academic Purposes*, *8*, 207–223.

Suk, N. (2017). The effects of extensive reading on reading comprehension, reading rate, and vocabulary acquisition. *Reading Research Quarterly*, *52*(1), 73–89.

Swain, M. (2001). Integrating language and content teaching through collaborative tasks. *Canadian Modern Language Review*, *58*(1), 44–63.

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Swales, J. M., & Feak, C. (2010). From text to task: Putting research on abstracts to work. In M. Ruiz-Garrido, J. C. Palmer-Silveira, & I. Fortanet-Gomez (Eds.), *Current trends in English for professional and academic purposes* (pp. 167–180). Rodopi.

The University of Manchester. (2020). Academic Phrasebank. Retrieved from http://www.phrasebank.manchester.ac.uk/

Therova, D. (2020a). Review of academic word lists. *TESL-EJ*, *24*(1), 1–15.

Therova, D. (2020b). General word lists: Overview and evaluation. *Vocabulary Learning and Instruction*, *9*(1), 51–61.

Thompson, P. (2005). Points of focus and position: Intertextual reference in PhD theses. *Journal of English for Academic Purposes*, *4*(4), 307–323.

Thompson, P. (2015a). The potential for corpus enhanced learning on the IFP. *InForm*, (15), 4–5.

Thompson, P. (2015b). Changing the bases for academic word lists. In P. Thompson & G. Diani (Eds.), *English for Academic Purposes: Approaches and implications* (pp. 317–342). Cambridge Scholars Publishing.

Thompson, P. (2016). Genre approaches to theses and dissertations. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for Academic Purposes* (pp. 379–391). Routledge.

Todd, R. W. (2017). An opaque engineering word list: Which words should a teacher focus on? *English for Specific Purposes*, *45*, 31–39.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. John Benjamins.

Tongpoon-Patanasorn, A. (2018). Developing a frequent technical words list for finance: A hybrid approach. *English for Specific Purposes*, *51*, 45–54.

Townsend, D., & Kiernan, D. (2015). Selecting academic vocabulary words worth learning. *Reading Teacher*, *69*(1), 113–118.

Tribble, C., & Wingate, U. (2013). From text to corpus - A genre-based approach to academic literacy instruction. *System*, *41*(2), 307–321.

Tshotsho, B. (2014). Assessing students' academic writing using systemic functional linguistics at a university in South Africa. *International Journal of Educational Sciences*, *6*(3), 425–433.

Tuck, J. (2015). 'Doing something that's really important': Meaningful engagement as a resource for teachers' transformative work with student writers in the disciplines. In T. Lillis, K. Harrington, M. R. Lea, & S. Mitchell (Eds.), *Working with academic literacies: Case studies towards transformative practice* (pp. 195–204). Parlor Press.

Turnitin. (n.d.). Retrieved from https://www.turnitin.com/

Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *Journal of English for Academic Purposes*, *12*(4), 248–263.

Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, *24*(1), 56–89.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, J., Liang, S., & Ge, G. (2008). Establishment of a Medical Academic Word List. *English for Specific Purposes*, *27*, 442–458.

Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes*, *28*, 170–182.

Watson, L., & Edwards, E. (2016). Making sense of global society: Teaching key conceptual vocabularies in the foundation year classroom. *InForm*, (16), 14–15.

Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assessment and Evaluation in Higher Education*, *31*(3), 379–394.

West, M. (1953). *A general service list of English words*. Longman.

Wingate, U. (2010). The impact of formative feedback on the development of academic writing. *Assessment and Evaluation in Higher Education*, *35*(5), 519–533.

Wingate, U. (2018). Academic literacy across the curriculum: Towards a collaborative instructional approach. *Language Teaching*, *51*(3), 349–364.

Wingate, U. (2019). Achieving transformation through collaboration: the role of academic literacies. *Journal of Learning Development in Higher Education*, (15), 1–9.

Wingate, U., & Tribble, C. (2012). The best of both worlds? Towards an English for Academic Purposes/Academic Literacies writing pedagogy. *Studies in Higher Education*, *37*(4), 481–495.

Woodward-Kron, R. (2005). The role of genre and embedded genres in tertiary students´ writing. *Prospect*, *20*(3), 24–44.

Wrigglesworth, J. (2019). Pedagogical applications of academic literacies theory: a reflection and case study. *Journal of Learning Development in Higher Education*, (15), 1–22.

Xudong, D., Cheng, L. K., Varaprasad, C., & Leng, L. M. (2010). Academic writing development of ESL/EFL graduate students in NUS. *Reflections on English Language Teaching*, *9*(2), 119–138.

Xue, G., & Nation, I. S. P. (1984). A University Word List. *Language Learning and Communication*, *3*(2), 215–229.

Yakhontova, T. (2002). 'Selling' or 'telling'? The issue of cultural variation in research genres. In J. Flowerdew (Ed.), *Academic discourse* (pp. 216–232). Longman.

Yang, H. (1986). A new technique for identifying scientific/technical terms and describing science texts. *Literacy and Linguistic Computing*, *1*, 93–103.

Yang, M. N. (2015). A nursing academic word list. *English for Specific Purposes*, *37*, 27–38.

Zhang, Z. (2011). Chinese international students' academic writing: An academic literacies perspective, *13*(1), 39–59.

Zimmerman, C. B. (1997). Do reading and interactive vocabulary instruction make a difference? An empirical study. *TESOL Quarterly*, *31*(1), 121–140.

Appendix

Appendix 1: Ethics approval

Human Research Ethics Committee (HREC)

The Open University

From [redacted]
The Open University Human Research Ethics Committee

Email Extension [redacted]

To Dana Therova

Project title: Linguistic complexity in foundation students' assessed writing at UK universities

HREC ref HREC/2694/Therova

Memorandum

Date application submitted: 19/10/2017
Date of HREC response: 03/11/2017

This memorandum is to confirm that the research protocol for the above-named research project, as submitted for ethics review, has been given a favourable opinion by the Open University Human Research Ethics Committee.

Please note the following:

1. You are responsible for notifying the HREC immediately of any information received by you, or of which you become aware which would cast doubt on, or alter, any information contained in the original application, or a later amendment which would raise questions about the safety and/or continued conduct of the research.

2. It is essential that any proposed amendments to the research are sent to the HREC for review so they can be recorded and where required, a favourable opinion given prior to any changes being implemented (except only in cases of emergency when the welfare of the participant or researcher is or may be effected).

3. Please include your HREC reference number in any documents or correspondence. It is essential that it is included in any publicity related to your research, e.g. when seeking participants or advertising your research so it is clear that it has been reviewed by HREC and adheres to OU ethics review processes.

4. You are authorised to present this memorandum to outside bodies such as NHS Research Ethics Committees in support of any application for future research clearance. Also, where there is an external ethics review, a copy of the application and outcome should be sent to the HREC.

5. OU research ethics review procedures are fully compliant with the majority of grant awarding bodies and where they exist, their frameworks for research ethics.

6. At the end of your project, you are required to assess your research for ethics related issues and/or major changes. Where these have occurred you will need to provide the Committee with a HREC final report to reflect how these were dealt with using the final

www.open.ac.uk/research/ethics/ November 2017

Appendix 2: Consent form

Name of Researcher:	Dana Therova
Researcher's address:	School of Languages and Applied Linguistics, The Open University, Stuart Hall Building, Level 3, Milton Keynes, Buckinghamshire, England, MK7 6AA
E-mail address:	dana.therova@open.ac.uk
Telephone:	[redacted]
Title of study:	Linguistic complexity in foundation students' assessed writing at UK universities

Student name: **Student number:**

Nationality: **Age:**

By signing this form, I confirm that:

- I have been provided with an Information Sheet about the above study and I have read and understood the information about the study.
- I have been given the opportunity to contact the research to ask questions about the study and my participation.
- I understand that my participation is voluntary and that I can withdraw at any time, but no later than 30.4.2018, without giving reason.
- I understand that my withdrawal from the study would not affect my future treatment at the University.
- I understand that all information about me will be treated in strict confidence (i.e. anonymously) and that I will not be named in any written or oral work arising from this study.
- I understand that the study will involve the use of my written assignments submitted to the University.
- If I agree to be interviewed, I agree to the interview to be audio and/or video recorded and I understand that any audio or video material of me will be used solely for research purposes. I also understand that the interview will take place via Skype or on the University's premises and will not last longer than 45 minutes.
- The archiving and sharing of the data in publications and at conferences has been explained to me in the Information Sheet and I agree to the use of anonymised data in publications and at conferences.
- I understand that other researchers might have access to this data only if they agree to preserve the confidentiality of the data and if they agree to the terms specified in this form.
- I understand that the progress of this study will be discussed with others at the Open University and that this will also be done anonymously.
- By signing this form, I freely give my consent to participate in this study and confirm that I have been given a copy of this form for my own information.

Signature: **Date:**

Appendix 3: Information sheet

Name of Researcher:	Dana Therova
Researcher's address:	School of Languages and Applied Linguistics, The Open University, Stuart Hall Building, Level 3, Milton Keynes, Buckinghamshire, England, MK7 6AA
E-mail address:	dana.therova@open.ac.uk
Telephone:	[redacted]

Title of study: Linguistic complexity in foundation students' assessed writing at UK universities

Introduction:
I would like to invite you to participate in this research study which aims to investigate foundation students' assessed writing with a particular focus on linguistic features such as academic vocabulary and syntactic structures.
Before you decide whether to take part in my study, please read the following information which explains what participation in this study would involve.

What is the purpose of this study?
This project is part of my PhD study which aims to investigate features of foundation students' writing with the aim to better understand academic literacies of foundation students from various educational and linguistic backgrounds. The practical implications of this study lie in informing foundation degree programmes which focus on the delivery of academic writing as well as relevant university provisions offering academic language support.

What are the aims of this study?
This study aims to investigate lexical and syntactic features displayed in written assignments of foundation students at UK universities and how they change over the course of one academic year. In addition, this study aims to establish what contributes to this development.

Why have I been invited?
You have been invited to participate in my study as part of your cohort as the varied cultural, linguistic and educational background of your cohort forms a suitable sample representing foundation students at UK universities.

Do I have to take part?
Participation in this study is entirely voluntary and so you are not obliged to take part in it. If you do not wish to take part, you do not have to give a reason and you will not be contacted again. Similarly, if you do agree to participate, you are free to withdraw at any time during the project if you change your mind. However, given the limited time available to collect and analyse the relevant data, you would have to inform me by 30.4.2018 if you wished to withdraw from the study.

What happens if I withdraw from the study?
If you decide that you do not wish to continue to participate in the study, you can request to have all data collected from you destroyed. Any data collected from you would thus not be used in the study in any form.

What will I have to do if I agree to take part?
Towards the end of the academic year I would like to interview a few students. With those who volunteer to participate in an interview we will arrange a convenient time to meet at the University or via Skype and there will an interview with myself during which I will ask you questions relating to your academic writing. The interview is expected to last no longer than 45 minutes and is going to be a one-off event. The interview will also be audio recorded which will enable me to transcribe it at a later stage. It is important for you to know that any information provided during the interview will be de-identified/anonymous i.e. your name will not appear anywhere. Your responses to my questions will be used for the purpose of this study only.

Will I have to do anything else?
In addition to interview data, I will also ask for your permission to use your written assignments that you have submitted to the University as part of your course requirements which will enable to me do a linguistic analysis of your language use. Apart from that, you will not be required to do anything else.

How will the data be used?
The data will be obtained for research purposes and will remain confidential and anonymous. This means that your name will not appear in any work produced on the basis of the data collected. However, the data that I collect will be discussed with my project supervisors at the Open University. This will also be done anonymously.

Will I have access to the results of the study?
On completion of my study, which is anticipated by the end 2021, I will produce a summary of the findings which I will be happy to share with all students who have participated in my study. If you are interested in the results, please get in contact with me using the contact details provided above.

Are there any risks associated with participation?
Since the data will be in the form of your written assignments and interview data, there are no anticipated risks or physical harm that this study could pose for you. As explained above, all data will be confidential and anonymous and will in no way affect your future treatment at the University. There will be no link between my study and your assessment results. The researcher will not be marking the assignments and the data collected will be entirely separate from marks allocated. In other words, the research does not involve any assessment of the standards of your written assignments, and the findings of the analysis will not affect in any way the assignment assessment.

What if I have issues or enquiries which cannot be satisfactorily resolved with the researcher?
If you feel that you need to contact another person regarding any matters or concerns relating to any aspects of the research project which cannot be satisfactorily resolved with me, please feel free to contact my lead research supervisor Prithvi Shrestha by emailing him at prithvi.shrestha@open.ac.uk.

What happens next?
If you are happy to take part in the study you are asked to complete the Consent Form supplied to you with this Information Sheet and return it to me. If you decide you do not wish to participate in the study, no response is required and no further contact will be made in relation to this study.

Thank you for taking the time to read this Information Sheet. If you have any further questions, please do not hesitate to approach me and I will be more than happy to address any questions or concerns that you may have.

Appendix 4: Overview of participants

UK (2014-2017)				
Total	**Nationalities/L1s**	**Ages**	**Gender**	**Intended UG programme**
25 students	17 nationalities 12 L1s	18-23	Male: 15 Female: 10	Business courses x12 Computer courses x5 Law x3 Psychology x2 Creative writing and media x1 Performing arts x1 Social sciences x1
UAE (2017-2018)				
Total	**Nationalities/L1s**	**Ages**	**Gender**	**Intended UG programme**
113 students	40 nationalities 38 L1s	16-25	Male: 43 Female: 70	Business courses x45 Psychology x19 Law x15 Creative writing and media x13 Computer courses x11 Not stated x6 Cinematics and games x3 Art and design x1
Mauritius (2017-2018)				
Total	**Nationalities/L1s**	**Ages**	**Gender**	**Intended UG programme**
55 students	14 nationalities 19 L1s	17-26	Male: 25 Female: 30	Law x17 Business courses x14 Computer courses x12 Psychology x6 Creative writing and media x4 Social sciences x1 Natural sciences x1
Overview (all students)				
Total	**Nationalities/L1s**	**Ages**	**Gender**	**Intended UG programme**
193 students	54 nationalities 55 L1s	16-26	Male: 83 Female: 110	Business courses x71 Law x35 Computer courses x28 Psychology x27 Creative media and writing x18 Other x8 Not stated x6

Nationalities*					
American	Angolan	Austrian	Bahraini	Bangladeshi	Bermudian
Brazilian	British	Burundian	Canadian	Chinese	Congolese
Cypriot	Egyptian	Emirati	Filipino	French	German
Ghanaian	Indian	Iranian	Italian	Jamaican	Japanese
Jordanian	Kazakhstani	Kenyan	Lebanese	Macedonian	Malagasy
Malawian	Mauritian	Moroccan	Nigerian	Pakistani	Philippines
Portuguese	Russian	Saudi Arabian	Seychellois	South African	South Korean
Sri Lankan	Sudanese	Syrian	Tanzanian	Turkish	Turkmen
Ugandan	Uzbekistan	Venezuelan	Vietnamese	Zambian	Zimbabwean

L1s*					
Afrikaans	Akan	Arabic	Bangla	Chaga	Chichewa
Chinese	Creole	Eleme	English	Farsi	Filipino
French	Ga	German	Gujarati	Hausa	Hindi
Igbo	Indian	Italian	Japanese	Kannada	Kazakh
Kikuyu	Kirundi	Kiswahili	Konkan	Korean	Kutchi
Lou	Macedonian	Malayalam	Marathi	Mauritian	Nigerian
Pashto	Portuguese	Punjabi	Russian	Saraiki	Shona
Sinhalese	Somali	Spanish	Swahili	Syrian	Tagalog
Tamil	Telugu	Turkish	Turkmen	Urdu	Uzbek
Vietnamese					

* information provided here is verbatim as stated by participants

Appendix 5: Interview schedule

Themes and Questions

1) Reasons for studying at a UK university

Q: Why did you choose to study at a UK university?

2) Prior experience of learning English with the focus on writing / academic writing

Q: Can you describe your prior experience of learning English with the focus on writing / academic writing i.e. was the focus on the writing process or the final product of your writing? In other words, would you be given any guidance as to how to approach writing tasks (writing process) or were you just given feedback on what you produced (writing product)?
Q: What different text types did you have to complete?
Q: Were most of the tasks formative or summative?
Q: Was the writing you had to do exam-driven i.e. focusing specifically on a completion of particular exam tasks such as the writing section in the IELTS test?
Q: What sort of feedback did you receive on your writing; what did it mainly relate to? (*prompts, if needed: grammar, vocabulary, text structure)*

3) IELTS

Q: How long before arriving in the UK / enrolling on the IFP did you take the IELTS test?
Q: How long did it take you to prepare for the test?
Q: Which part of the test (reading, writing, listening, speaking) did you find most difficult and why?
Q: How helpful did you find the IELTS test in preparation for academic writing at a UK university / for your first essay on the IFP?
Q: Has it influenced the way you approached written academic tasks on this programme / your first essay? If so, how?
Q: In terms of the development of your vocabulary, did you find that the IELTS had prepared you sufficiently for university study / IFP?
Q: In terms of the development of your grammar / sentence structure, how helpful did you find the IELTS?
Q: What IELTS score did you get for the writing part and overall?

4) Perceived differences between UK academic writing and prior experience

Q: What would you say are the main differences between UK academic writing and your prior experience?
(prompts, if needed: in terms of language use, level of criticality, research/reading-based, content, structure)

5) Students' perceptions of what constitutes good academic writing

Q: What, in your opinion, constitutes good academic writing at university?

6) Perceived difficulties with the writing process

Q: What have you found particularly difficult in terms of the process of completing written academic tasks such as essays and reports?
(*prompts, if needed: understanding of task and requirements, the amount of reading/research, language use/register, grammar, use of appropriate vocabulary, structuring arguments, organisation of ideas, criticality, referencing, paraphrasing, summarising, use of signposting and linking devices)*

7) Writing progression

Q: How do you think your writing has developed over the course of the academic year?
Q: What do you think has contributed to this?
(*prompts, if needed: taught sessions, independent learning, reading)*

8) Vocabulary development

Q: (How) do you think your academic vocabulary used in written assignments has developed over the course of the academic year?
Q: What do you think has contributed to this?
(prompts, if needed: taught sessions, independent learning, reading)

9) Cognitive processes employed by students

Q: Can you describe to me how you approach written academic tasks from title to completion i.e. what do you do first, next etc.? How long did each of these stages take you and why? Did you do many drafts? What was particularly helpful and why?

10) Writing strategies employed by students

Q: Can you describe to me how you go about / went about the actual writing of your assignments e.g. what are some of the strategies that you use such as paraphrasing, editing, translating, proof-reading etc.?

11) Perceived progress and contributors

Q: What have you learned from your feedback on your assignments? How are you going to approach future written academic tasks, based on your feedback? Are you going to do anything differently? If so, what and how?
Q: Do you think your academic writing has changed since you took IELTS / since the start of the IFP? If so, how and what contributed to it?
Q: Have you found the teaching materials useful? If so, what was particularly useful?
Q: What are the most important things that you have learned on the IFP in terms of academic writing?

12) Literacy beyond the IFP

Q: What other writing do you do outside of the university?
Q: How much reading and writing do you do in your language and in English?

13) Other

Q: Is there anything else that you would like to tell me about your experience with academic writing so far?
Q: Is there anything else that you would like to add that we have not discussed?

Appendix 6: Online survey

Dear IFP Student,
Thank you for agreeing to participate in my study which explores IFP students' academic writing. In addition to your written assignments, I would like to gain insights into your experience with academic writing. Therefore, I would very much appreciate you taking the time to complete this survey. I would be very grateful if you could provide as much information as possible. Thank you in advance.

Your name (optional):

1. Why did you choose to study at a UK university?

2. What different types of writing did you experience prior to enrolling on the IFP?

- essays
- reports
- short paragraphs
- other…

3. What was the purpose of the writing you did prior to the IFP?

- to prepare for an exam (e.g. IELTS)
- for assessment
- to receive feedback (formative - not assessed)
- to practise / improve writing
- other…

4. What type of feedback did you receive on your writing style prior to the IFP?

- feedback on grammar
- feedback on vocabulary
- feedback on text organisation
- feedback on the content
- other…

5. How long before enrolling on the IFP did you take the IELTS test?

- up to 3 months
- 4-8 months
- 9-12 months
- more than 1 year

6. How long did it take you to prepare for the IELTS test?

- up to 3 months
- 4-8 months
- 9-12 months
- more than 1 year

7. Which part of the IELTS test did you find most difficult?

- reading
- writing
- speaking
- listening

8. In terms of preparation for your written assignments (essays and reports) on the IFP, how helpful was the IELTS?

0 = not helpful at all 10 = most helpful

0 1 2 3 4 5 6 7 8 9 10

9. In terms of vocabulary used in your essays and reports on the IFP, how helpful was the IELTS?

0 = not helpful at all 10 = most helpful

0 1 2 3 4 5 6 7 8 9 10

10. In terms of grammar/sentence structure you used in your essays and reports on the IFP, how helpful was the IELTS?

0 = not helpful at all 10 = most helpful

0 1 2 3 4 5 6 7 8 9 10

11. What did you find particularly helpful in completing your essays and reports on the IFP? (tick as many as you like)

- information provided in lectures, seminars and labs
- teaching materials
- seeing examples
- receiving feedback on drafts
- peer support and guidance
- independent study
- other…

12. What did you find difficult in terms of the process of completing your essays and reports on the IFP? (tick as many as you like)

- understanding the task requirements
- the amount or reading/research required
- use of appropriate vocabulary
- use of appropriate grammar and sentence structure
- use of correct punctuation
- use of articles (a, an, the)
- using own words
- structuring and organising ideas/arguments in paragraphs and/or whole text
- structure of reports
- other…

13. How many drafts do you typically do before submitting your essays and reports on the IFP?

- 0
- 1
- 2
- 3 or more

14. Which of the following writing strategies did you use when completing your essays and reports on the IFP? (tick as many as you like)

- paraphrasing (i.e. using own words)
- summarising (i.e. extracting main points)
- translating
- using direct quotes (i.e. using the wording of the source text)
- using vocabulary from texts
- using vocabulary from assignments briefs/instructions
- using vocabulary from teaching materials
- other…

15. List any new words you learned from the process of completing your essays and reports on the IFP. (list as many as you like)

16. How do you think your writing has changed during the IFP (e.g. in terms of grammar, sentence structure, vocabulary etc.)?

17. If you feel your writing has changed over the course of the IFP, what do you think has contributed to it? (tick as many as you like)

- teaching
- reading
- seeing examples in sources
- receiving feedback on drafts
- independent learning
- other…

18. Based on your experience on the IFP, what in your opinion constitutes good academic writing at university?

19. What would you say are the main differences between the academic writing you were required to use on the IFP and your prior experience?

20. What are the most important things you have learned about academic writing on the IFP?

21. What do you still find difficult in terms of completing written academic tasks such as essays and reports?

22. What do you feel you still need to continue to improve in terms of academic writing? How are you going to continue to improve?

23. Would you be willing to be interviewed about your experience with academic writing on the IFP?

- Yes
- No

If you are happy to be interviewed, please type in your name and email address below and I will get in contact with you to arrange a Skype interview.

Appendix 7: Academic vocabulary used by at least 50% of students

Appendix 7a: Problem Question PQ_BOS_48

Academic vocabulary used by at least 50% of students (PQ_BOS_48)									
Academic lemma	**Range**	**Raw freq.**	**Norm freq.**	**Function**	**Academic lemma**	**Range**	**Raw freq.**	**Norm freq.**	**Function**
review	48	79	9.76	TO-struc	*beneficial**	31	182	22.48	RO-desc
literature	48	70	8.65	TO-struc	*group*	31	102	12.6	RO-desc
*structure**	47	864	106.72	RO-top	*task*	31	72	8.89	RO-proc
*project**	46	498	61.51	RO-proc	*define*	31	63	7.78	TO-fram
introduction	46	52	6.42	TO-struc	*research**	30	117	14.45	RO-proc
control	45	283	34.96	RO-top	*include*	30	79	9.76	RO-proc
conclusion	45	70	8.65	TO-struc	*support*	30	78	9.63	RO-proc
*environment**	42	250	30.88	RO-top	*process*	30	66	8.15	RO-proc
environmental	42	198	24.46	RO-desc	*develop*	30	48	5.93	RO-proc
hierarchy	42	171	21.12	RO-top	*activity*	29	94	11.61	RO-proc
need	41	160	19.76	RO-proc	*achieve*	29	58	7.16	RO-proc
organization	40	354	43.73	RO-top	*compare**	29	55	6.79	RO-proc
sustainable	40	179	22.11	RO-desc	*contribute*	29	43	5.31	RO-proc
level	40	171	21.12	RO-desc	*however*	28	86	10.62	TO-trans
*both**	40	170	21	TO-fram	*development*	27	76	9.39	RO-proc
provide	40	147	18.16	RO-proc	*goal*	27	76	9.39	RO-desc
management	39	198	24.46	RO-top	*quality*	27	67	8.28	RO-desc
information	39	100	12.35	RO-desc	*value*	27	66	8.15	RO-top
reduce	38	126	15.56	RO-proc	*impact*	27	60	7.41	RO-proc
role	38	97	11.98	RO-des	*furthermore*	27	59	7.29	TO-trans
benefit	37	99	12.23	RO-proc	*culture*	26	119	14.7	RO-desc
base	37	78	9.63	RO-desc	*initiative*	26	87	10.75	RO-proc
important	37	76	9.39	RO-desc	*material*	26	86	10.62	RO-desc
social	36	161	19.89	RO-desc	*produce*	26	75	9.26	RO-proc
aim	36	102	12.6	RO-proc	*example*	26	66	8.15	TO-struc
improve	35	55	6.79	RO-proc	*low*	26	62	7.66	RO-desc
focus	34	87	10.75	RO-proc	*purpose*	26	40	4.94	TO-struc
type	34	85	10.5	RO-desc	*comparison*	26	31	3.83	RO-desc
sector	33	147	18.16	RO-desc	*system*	25	41	5.06	RO-desc
communication	33	92	11.36	RO-proc	*term**	25	38	4.69	TO-fram
society	33	83	10.25	RO-desc	*therefore*	24	62	7.66	TO-res
increase	33	70	8.65	RO-proc	*source*	24	53	6.55	RO-desc
resource	33	70	8.65	RO-desc	*technology*	24	46	5.68	RO-desc
product	32	173	21.37	RO-proc	*discuss*	24	45	5.56	TO-struc
waste	32	113	13.96	RO-proc	*general*	24	44	5.43	TO-fram
sustainability	32	112	13.83	RO-top	*global*	24	32	3.95	RO-desc
within	32	76	9.39	TO-fram					

*academic vocabulary item in the assignment brief

Appendix 7b: Problem Question PQ_ESP_73

Academic vocabulary used by at least 50% of students (PQ_ESP_73)									
Academic lemma	**Range**	**Raw freq.**	**Norm freq.**	**Function**	**Academic lemma**	**Range**	**Raw freq.**	**Norm freq.**	**Function**
*ethical**	73	934	78.55	RO-top	*provide*	52	127	10.68	RO-proc
*research**	72	688	57.86	RO-top	*present*	50	139	11.69	RO-desc
review	72	117	9.84	TO-struc	*act*	50	107	9	RO-proc
conclusion	72	114	9.59	TO-struc	*aim*	49	127	10.68	TO-proc
experiment	71	2172	182.66	RO-top	*individual*	49	117	9.84	RO-desc
study	71	1001	84.18	RO-desc	*subject*	48	273	22.96	RO-top
consent	70	385	32.38	RO-top	*human*	47	144	12.11	RO-desc
literature	70	84	7.06	TO-struc	*discuss*	47	103	8.66	TO-struc
conduct	69	357	30.02	RO-proc	*participate*	46	98	8.24	RO-proc
*ethic**	69	314	26.41	RO-top	*purpose*	45	122	10.26	TO-struc
introduction	69	74	6.22	TO-struc	*furthermore*	45	110	9.25	TO-trans
*social**	68	368	30.95	RO-top	*affect*	44	87	7.32	RO-proc
state	68	236	19.85	TO-struc	*group*	43	229	19.26	RO-desc
*guideline**	67	465	39.11	RO-top	*involve*	43	129	10.85	RO-proc
psychological	66	321	27	RO-top	*both*	42	74	6.22	TO-fram
solution	66	198	16.65	RO-desc	*society*	42	73	6.14	RO-desc
result	65	313	26.32	TO-res	*ensure*	41	90	7.57	RO-proc
researcher	64	351	29.52	RO-top	*mental*	40	96	8.07	RO-desc
*inform**	61	204	17.16	RO-proc	*example*	40	54	4.54	TO-struc
however	58	192	16.15	TO-trans	*moreover*	39	88	7.4	TO-trans
informed	58	168	14.13	RO-top	*method*	38	64	5.38	RO-proc
critique	58	110	9.25	RO-proc	*lack*	38	62	5.21	RO-desc
*role**	56	147	12.36	RO-desc	*code*	37	118	9.92	RO-top
protection	56	132	11.1	RO-top	*test*	37	93	7.82	RO-top
information	55	161	13.54	RO-desc	*observe*	37	79	6.64	RO-proc
effect	54	131	11.02	RO-proc	*important*	37	75	6.31	RO-desc
include	53	120	10.09	RO-proc	*condition*	37	58	4.88	RO-top
university	53	96	8.07	RO-loc					

*academic vocabulary item in the assignment brief

Appendix 7c: Problem Question PQ_ST_59

Academic vocabulary used by at least 50% of students (PQ_ST_59)									
Academic lemma	**Range**	**Raw freq.**	**Norm freq.**	**Function**	**Academic lemma**	**Range**	**Raw freq.**	**Norm freq.**	**Function**
*sustainable**	59	1489	134.56	RO-top	*study*	43	142	12.83	RO-desc
*environment**	59	427	38.59	RO-top	*both*	43	117	10.57	TO-fram
review	59	141	12.74	TO-struc	*organisation*	42	375	33.89	RO-top
literature	59	124	11.21	TO-struc	*challenge*	42	113	10.21	RO-desc
conclusion	58	97	8.77	TO-struc	*product*	42	92	8.31	RO-proc
impact	57	420	37.95	RO-proc	*discuss**	42	90	8.13	TO-struc
introduction	57	80	7.23	TO-struc	*finding*	42	83	7.5	TO-struc
information	55	237	21.42	RO-desc	*example*	40	109	9.85	TO-struc
environmental	55	192	17.35	RO-top	*develop*	40	98	8.86	RO-proc
abstract	55	58	5.24	TO-struc	*barrier**	39	141	12.74	RO-top
increase	54	266	24.04	RO-proc	*identify*	39	64	5.78	RO-proc
*research**	54	221	19.97	RO-proc	*aim*	38	90	8.13	RO-proc
economic	54	213	19.25	RO-desc	*achieve*	38	83	7.5	RO-proc
activity	54	183	16.54	RO-proc	*effect*	38	55	4.97	RO-proc
source	54	157	14.19	RO-desc	*group*	37	150	13.56	RO-desc
promote	53	297	26.84	RO-proc	*encourage*	37	82	7.41	RO-proc
*development**	53	294	26.57	RO-proc	*infrastructure*	37	81	7.32	RO-top
resource	53	158	14.28	RO-desc	*purpose*	37	58	5.24	TO-struc
*practice**	52	274	24.76	RO-proc	*experience*	37	54	4.88	RO-proc
sector	51	224	20.24	RO-des	*lack*	36	72	6.51	RO-desc
provide	51	184	16.63	RO-proc	*international*	36	70	6.33	RO-desc
improve	51	138	12.47	RO-proc	*moreover*	35	95	8.58	TO-trans
sustainability	50	309	27.92	RO-top	*furthermore*	34	71	6.42	TO-trans
contribute	50	101	9.13	RO-proc	*focus*	34	63	5.69	RO-proc
however	49	161	14.55	TO-trans	*waste*	33	105	9.49	RO-proc
social	49	148	13.37	RO-desc	*initiative*	33	97	8.77	RO-proc
employment	49	128	11.57	RO-top	*management*	33	88	7.95	RO-proc
result	49	124	11.21	TO-res	*ensure*	33	80	7.23	RO-proc
future	49	117	10.57	RO-desc	*change*	32	64	5.78	RO-proc
*project**	48	219	19.79	RO-proc	*society*	32	63	5.69	RO-desc
discussion	48	78	7.05	TO-struc	*involve*	31	87	7.86	RO-proc
negative	47	132	11.93	RO-desc	*state*	31	67	6.05	TO-struc
*support**	47	126	11.39	RO-proc	*base*	31	57	5.15	RO-desc
positive	47	115	10.39	RO-desc	*maintain*	31	54	4.88	RO-proc
reduce	46	179	16.18	RO-proc	*explore*	31	42	3.8	RO-proc
benefit	46	153	13.83	RO-proc	*population*	30	58	5.24	RO-desc
need	46	146	13.19	RO-proc	*therefore*	30	56	5.06	TO-res
culture	46	139	12.56	RO-desc	*article*	30	55	4.97	RO-desc
*include**	46	114	10.3	RO-proc	*heritage*	30	48	4.34	RO-desc
growth	46	93	8.4	RO-proc	*process*	30	47	4.25	RO-proc
recommendation	46	77	6.96	TO-struc	*organization**	29	205	18.53	RO-top
natural	45	165	14.91	RO-desc	*system*	29	77	6.96	RO-desc
cultural	45	146	13.19	RO-desc	*thus*	29	76	6.87	TO-res

importance	45	119	10.75	RO-desc	*conserve*	29	49	4.43	RO-top
*affect**	45	111	10.03	RO-proc	*protection*	29	47	4.25	RO-top
*important**	44	113	10.21	RO-desc	*nature*	29	44	3.98	RO-top
generate	44	103	9.31	RO-proc	*require*	29	43	3.89	RO-proc
method	44	102	9.22	RO-proc					

*academic vocabulary item in the assignment brief

Appendix 7d: Problem Question PQ_SI_13

Academic vocabulary used by at least 50% of students (PQ_SI_13)									
Academic lemma	**Range**	**Raw freq.**	**Norm freq.**	**Function**	**Academic lemma**	**Range**	**Raw freq.**	**Norm freq.**	**Function**
*initiative**	13	204	94.42	RO-top	*therefore*	7	30	13.89	TO-res
reduce	13	73	33.79	RO-proc	*method*	7	20	9.26	RO-proc
*barrier**	13	47	21.75	RO-desc	*system*	7	19	8.79	RO-desc
future	13	39	18.05	RO-desc	*type*	7	18	8.33	RO-desc
implementation	13	37	17.13	RO-proc	*furthermore*	7	16	7.41	TO-trans
literature	13	18	8.33	TO-struc	*growth*	7	12	5.55	RO-proc
review	13	18	8.33	TO-struc	*individual*	7	12	5.55	RO-desc
*introduction**	13	17	7.87	TO-struc	*global*	7	11	5.09	RO-desc
conclusion	13	15	6.94	TO-struc	*low*	7	11	5.09	RO-desc
*sustainable**	12	157	72.67	RO-top	*involve*	7	10	4.63	RO-proc
*benefit**	12	82	37.95	RO-proc	*important*	7	9	4.17	RO-desc
*development**	12	79	36.57	RO-proc	*rate*	7	9	4.17	RO-desc
produce	12	31	14.35	RO-proc	*term*	7	9	4.17	TO-fram
*environment**	11	62	28.7	RO-top	*essential*	7	8	3.7	RO-desc
*sustainability**	11	48	22.22	RO-top	*purpose*	7	8	3.7	TO-struc
resource	11	47	21.75	RO-desc	*product*	6	25	11.57	RO-proc
increase	11	42	19.44	RO-proc	*consumption*	6	21	9.72	RO-proc
environmental	11	37	17.13	RO-desc	*hence*	6	18	8.33	TO-res
economic	11	35	16.2	RO-desc	*generate*	6	17	7.87	RO-proc
natural	11	29	13.42	RO-desc	*encourage*	6	16	7.41	RO-proc
*include**	11	23	10.65	RO-proc	*various*	6	16	7.41	RO-desc
project	10	63	29.16	RO-proc	*material*	6	15	6.94	RO-desc
aim	10	25	11.57	RO-proc	*within**	6	15	6.94	TO-fram
present	10	21	9.72	RO-desc	*however*	6	14	6.48	TO-trans
state	10	16	7.41	TO-struc	*change*	6	13	6.02	RO-proc
process	9	34	15.74	RO-proc	*impact**	6	13	6.02	RO-proc
require	9	23	10.65	RO-proc	*production*	6	13	6.02	RO-proc
achieve	9	22	10.18	RO-proc	*discuss**	6	12	5.55	TO-struc
example	9	19	8.79	TO-struc	*solution*	6	12	5.55	RO-desc
source	9	19	8.79	RO-desc	*both*	6	11	5.09	TO-fram
waste	8	79	36.57	RO-proc	*goal*	6	11	5.09	RO-desc
promote	8	22	10.18	RO-proc	*study*	6	10	4.63	RO-desc
provide	8	20	9.26	RO-proc	*define*	6	9	4.17	TO-fram
social	8	18	8.33	RO-desc	*effect*	6	9	4.17	RO-proc
develop	8	16	7.41	RO-proc	*firstly*	6	9	4.17	TO-struc
form	8	15	6.94	RO-desc	*nature*	6	8	3.7	RO-top
improve	8	14	6.48	RO-proc	*research**	6	8	3.7	RO-proc
population	8	13	6.02	RO-desc	*available*	6	6	2.78	RO-desc
*beneficial**	8	11	5.09	RO-desc	*conclude*	6	6	2.78	TO-struc
technology	8	11	5.09	RO-desc	*requirement*	6	6	2.78	RO-desc
usage	8	11	5.09	RO-proc					

*academic vocabulary item in the assignment brief

Appendix 7e: Research Report RR_193

Academic vocabulary used by at least 50% of students (RR_193)									
Type	**Range**	**Raw freq.**	**Norm freq.**	**Function**	**Type**	**Range**	**Raw freq.**	**Norm freq.**	**Function**
*review**	193	457	13.1	TO-struc	*solution*	120	285	8.17	RO-desc
*literature**	193	442	12.67	TO-struc	*change*	118	354	10.15	RO-proc
*barrier**	191	3242	92.95	RO-top	*state*	116	389	11.15	TO-struc
*research**	189	1337	38.33	RO-proc	*group**	116	371	10.64	RO-desc
introduction	188	234	6.71	TO-struc	*support*	116	343	9.83	RO-proc
*sustainable**	183	2384	68.35	RO-top	*base**	116	267	7.66	RO-desc
*conclusion**	182	323	9.26	TO-struc	*discuss*	115	259	7.43	TO-struc
abstract	179	184	5.28	TO-struc	*knowledge*	114	367	10.52	RO-desc
*study**	173	942	27.01	RO-desc	*benefit*	113	276	7.91	RO-proc
result	173	784	22.48	TO-struc	*resource*	113	268	7.68	RO-desc
*development**	164	1354	38.82	RO-top	*discussion*	113	145	4.16	TO-struc
*initiative**	159	1459	41.83	RO-proc	*focus*	112	231	6.62	RO-proc
lack	155	694	19.9	RO-desc	*finding*	111	245	7.02	TO-struc
*sustainability**	153	939	26.92	RO-top	*project**	110	486	13.93	RO-proc
*primary**	151	506	14.51	RO-desc	*therefore*	110	325	9.32	TO-res
*survey**	149	1337	38.33	RO-proc	*affect**	110	304	8.72	RO-proc
social	149	766	21.96	RO-desc	*above*	110	291	8.34	TO-struc
environment	149	608	17.43	RO-top	*present**	109	234	6.71	RO-desc
increase	143	447	12.82	RO-proc	*important*	108	272	7.8	RO-desc
however	142	552	15.83	TO-trans	*objective*	108	192	5.5	TO-struc
*provide**	141	432	12.39	RO-proc	*general**	106	224	6.42	TO-fram
*strategy**	138	368	10.55	TO-struc	*impact*	104	255	7.31	RO-proc
method	135	688	19.73	RO-proc	*furthermore*	103	243	6.97	TO-trans
economic	133	446	12.79	RO-desc	*system*	102	382	10.95	RO-desc
information	133	386	11.07	RO-desc	*achieve*	102	262	7.51	RO-proc
reduce	131	475	13.62	RO-proc	*challenge*	101	256	7.34	RO-desc
analysis	131	276	7.91	RO-proc	*graph**	99	261	7.48	TO-struc
*conduct**	130	358	10.26	RO-proc	*concern*	98	227	6.51	RO-desc
environmental	129	379	10.87	RO-desc	*factor*	98	217	6.22	RO-desc
identify	128	360	10.32	RO-proc	*process*	98	197	5.65	RO-proc
*include**	128	334	9.58	RO-proc	*both*	98	196	5.62	TO-fram
figure	125	1125	32.26	TO-struc	*develop*	97	254	7.28	RO-proc
population	124	461	13.22	RO-desc	*analyse**	97	222	6.37	RO-proc

*academic vocabulary item in the assignment brief

Appendix 7f: Discussion Essay E2_SMC_42

Academic vocabulary used by at least 50% of students (E2_SMC_42)									
Academic lemma	**Range**	**Raw freq.**	**Norm freq.**	**Function**	**Academic lemma**	**Range**	**Raw freq.**	**Norm freq.**	**Function**
*social**	42	1402	251.19	RO-top	*include**	28	53	9.5	TO-fram
information	40	237	42.46	RO-top	*form*	27	58	10.39	RO-desc
user	37	162	29.02	RO-top	*application*	26	45	8.06	RO-top
example	36	111	19.89	TO-struc	*group*	25	80	14.33	RO-desc
however	36	91	16.3	TO-trans	*content*	25	43	7.7	RO-desc
conclude	36	50	8.96	TO-struc	*society*	24	59	10.57	RO-desc
individual	35	117	20.96	RO-desc	*communicate*	24	46	8.24	RO-proc
*encourage**	35	105	18.81	RO-top	*enable*	24	44	7.88	RO-proc
provide	35	84	15.05	RO-proc	*discuss**	24	41	7.35	TO-struc
networking	34	152	27.23	RO-top	*result*	22	54	9.67	TO-res
report	33	94	16.84	RO-proc	*benefit*	22	46	8.24	RO-desc
act	32	89	15.95	RO-proc	*identify**	22	42	7.52	RO-proc
communication	32	69	12.36	RO-proc	*evidence**	22	37	6.63	RO-top
network	31	93	16.66	RO-top	*link*	22	36	6.45	RO-top
activity	31	85	15.23	RO-proc	*purpose*	22	36	6.45	RO-desc
furthermore	31	60	10.75	TO-trans	*conclusion**	22	24	4.3	TO-struc
increase	30	89	15.95	RO-proc	*thus*	21	47	8.42	TO-res
access	30	70	12.54	RO-proc	*negative**	21	42	7.52	RO-desc
define	30	59	10.57	TO-fram	*research**	21	40	7.17	RO-proc
tool	28	70	12.54	RO-desc	*positive**	21	37	6.63	RO-desc
*state**	28	64	11.47	TO-struc	*change*	21	33	5.91	RO-proc
technology	28	56	10.03	RO-top	*common*	21	28	5.02	RO-desc
moreover	28	55	9.85	TO-trans					

*academic vocabulary item in the assignment brief

Appendix 7g: Discussion Essay E2_SMA_38

Academic vocabulary used by at least 50% of students (E2_SMA_38)									
Academic lemma	**Range**	**Raw freq.**	**Norm freq.**	**Function**	**Academic lemma**	**Range**	**Raw freq.**	**Norm freq.**	**Function**
*social**	38	1264	246.13	RO-top	*benefit*	24	49	9.54	RO-desc
use	38	528	102.81	RO-proc	*therefore*	23	72	14.02	TO-res
*university**	37	299	58.22	RO-loc	*tend*	23	56	10.9	RO-proc
information	37	222	43.23	RO-top	*technology*	23	52	10.13	RO-top
*research**	37	147	28.62	RO-proc	*beneficial**	23	50	9.74	RO-desc
study	36	157	30.57	RO-top	*interaction*	23	43	8.37	RO-proc
communication	36	128	24.92	RO-proc	*enable*	23	32	6.23	RO-proc
networking	35	279	54.33	RO-top	*state*	22	65	12.66	TO-struc
*academic**	34	209	40.7	RO-top	*active*	22	54	10.52	RO-desc
group	34	88	17.14	RO-desc	*improve*	22	52	10.13	RO-proc
communicate	33	84	16.36	RO-proc	*both**	22	45	8.76	TO-fram
develop	32	74	14.41	RO-proc	*disadvantage*	21	45	8.76	RO-desc
*positive**	31	75	14.6	RO-desc	*promote*	21	43	8.37	RO-proc
conclude	31	56	10.9	TO-struc	*lack*	21	42	8.18	RO-desc
tool	30	76	14.8	RO-desc	*interact*	21	37	7.2	RO-proc
knowledge	30	56	10.9	RO-top	*performance*	20	58	11.29	RO-proc
*impact**	29	96	18.69	RO-proc	*thus*	20	54	10.52	TO-res
*negative**	29	74	14.41	RO-desc	*individual*	20	53	10.32	RO-desc
access	29	65	12.66	RO-proc	*moreover*	20	48	9.35	TO-trans
*experience**	28	112	21.81	RO-top	*process*	20	40	7.79	RO-proc
*network**	28	96	18.69	RO-top	*engage*	20	39	7.59	RO-proc
level	28	91	17.72	RO-top	*support**	20	38	7.4	RO-proc
however	28	71	13.83	TO-trans	*furthermore*	20	36	7.01	TO-trans
provide	28	70	13.63	RO-proc	*conclusion**	20	24	4.67	TO-struc
user	28	68	13.24	RO-top	*application*	19	49	9.54	RO-top
enhance	28	54	10.52	RO-proc	*content*	19	47	9.15	RO-desc
*purpose**	27	93	18.11	RO-desc	*aspect*	19	45	8.76	RO-desc
increase	27	58	11.29	RO-proc	*discussion*	19	45	8.76	RO-proc
result	26	73	14.21	TO-res	*lecturer*	19	45	8.76	RO-top
example	26	71	13.83	TO-struc	*depression*	19	43	8.37	RO-desc
affect	26	61	11.88	RO-proc	*usage**	19	40	7.79	RO-proc
*discuss**	26	53	10.32	TO-struc	*task*	19	38	7.4	RO-top
activity	25	51	9.93	RO-proc	*control*	19	32	6.23	RO-proc
important	25	50	9.74	RO-desc	*mean*	19	30	5.84	RO-desc
*effect**	24	51	9.93	RO-top	*mental*	19	30	5.84	RO-desc

*academic vocabulary item in the assignment brief

Appendix 7h: Discussion Essay E2_SS_113

Academic vocabulary used by at least 50% of students (E2_SS_113)									
Academic lemma	**Range**	**Raw freq.**	**Norm freq.**	**Function**	**Academic lemma**	**Range**	**Raw freq.**	**Norm freq.**	**Function**
*society**	110	1082	81.36	RO-top	*example*	69	154	11.58	TO-struc
information	104	526	39.55	RO-top	*argue*	68	148	11.13	TO-struc
monitor	97	352	26.47	RO-top	*increase*	67	176	13.23	RO-proc
however	97	329	24.74	TO-trans	*reduce*	66	167	12.56	RO-proc
individual	90	446	33.54	RO-desc	*firstly*	66	76	5.71	TO-struc
*benefit**	88	306	23.01	RO-desc	*evidence**	65	188	14.14	RO-top
*conclusion**	84	101	7.59	TO-struc	*social*	64	193	14.51	RO-desc
technology	81	283	21.28	RO-top	*rate*	64	181	13.61	RO-desc
system	80	387	29.1	RO-top	*human*	64	138	10.38	RO-desc
state	78	226	16.99	TO-struc	*furthermore*	63	120	9.02	TO-trans
*discuss**	78	135	10.15	TO-struc	*control*	62	156	11.73	RO-proc
therefore	74	202	15.19	TO-res	*identify**	59	113	8.5	RO-proc
activity	73	211	15.87	RO-proc	*act*	58	119	8.95	RO-proc
provide	72	203	15.26	RO-proc	*moreover*	56	89	6.69	TO-trans
purpose	72	159	11.96	RO-desc					

*academic vocabulary item in the assignment brief